UNIVERSITY OF CALIFORNIA SAN DIEGO

How the Structure of Choices Influences Consumer Decisions and Experiences

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

in

Management

by

Kristen Elizabeth Duke

Committee in charge:

Professor On Amir, Chair
Professor Ayelet Gneezy
Professor Jonathan Levav
Professor Wendy Liu
Professor Piotr Winkielman

2019

The Dissertation of Kristen Elizabeth Duke is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

2019

DEDICATION

I dedicate this dissertation to my loving and supportive parents, Cindy and John, and my partner for life, Tyler.

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Table of Contents v

List of Figures and Tables vii

List of Supplemental Files viii

Acknowledgements ix

Vita xiv

Abstract of the Dissertation xv

Chapter 1 Guilt Dynamics: Consequences of Temporally Separating Decisions and Actions 1
- Abstract 2
- Introduction 3
- Guilt and the Role of Personal Responsibility 4
- Two Sources of Guilt 6
- The Temporal Dynamics of Guilt 7
- Guilt Dynamics Framework 9
- The Behavioral Consequences of Guilt 12
- Overview of the Experiments 13
- Pilot Study 1 14
- Pilot Study 2 18
- Experiment 1 22
- Experiment 2 27
- Experiment 3 32
- Experiment 4 38
- General Discussion 43
- Acknowledgements 53
- Figures 54
- Appendix 60
- References 71

Chapter 2 The Quantity Integration Effect: When Integrating Purchase and Quantity Decisions Increases Sales 77
- Abstract 78
- Introduction 79

Theory 80
Data Overview 86
Experiment 1A: Sample Lab Experiment, Time-Controlled 87
Experiment 1B: Sample Lab Experiment, Website Navigation 89
Experiments 2A—2C: Support for a Novel Mechanism, Closure 93
Pooled Analysis 98
Field Experiment 102
Assessing the Role of Different Features and Mechanisms 106
General Discussion 110
Acknowledgments 117
Figures 118
Tables 123
Appendix 124
References 149

Chapter 3 How Incentive Framing Can Harness the Power of Social Norms 155
Abstract 156
Introduction 157
Overview of Studies 169
Study 1 170
Study 2 176
Study 3 181
Study 4 189
General Discussion 196
Acknowledgements 206
Figures 207
Appendix 211
References 225

LIST OF FIGURES AND TABLES

Figure 1.1 Dynamics of guilt with and without decision-enactment gaps. The left 54

Figure 1.2 Comparing predicted overall guilt with and without a decision-enactment 54

Figure 1.3 General overview of experimental timelines, Experiments 1—4..... 55

Figure 1.4 Current or anticipated guilt in each condition, Pilot Study 1. Error bars 55

Figure 1.5: Self-reported emotions, Experiment 1. Errors bars represent ±1 SE..... 56

Figure 1.6: The decay of decision guilt, Experiment 2. In the left panel, the left and..... 57

Figure 1.7: Covariate-adjusted amount of food consumed, Experiment 3. Outliers..... 58

Figure 1.8 Covariate-adjusted amount donated, Experiment 4. The covariates were..... 59

Figure 2.1: Schematic of QS and QI selling formats. This image depicts a typical 118

Figure 2.2: Theoretical implementation of QS and QI selling formats.: In a 119

Figure 2.3: Practical implementation of QS and QI selling formats..... 119

Figure 2.4: Choice screen for each selling format, Experiment 1B..... 120

Figure 2.5: Difference in purchase rate between selling formats across split..... 121

Figure 2.6: Schematic of QS finder interface..... 122

Figure 2.7: Schematic of QI finder interface 122

Table 2.1: Logistic regressions and linear probability models predicting purchase 123

Figure 3.1: The effects of incentive framing on perceived injunctive norms 207

Figure 3.2: The effects of incentive framing on perceived injunctive norms 208

Figure 3.3: The effects of incentive framing on downstream intention to bring 209

Figure 3.4: Regression-predicted intention to dress up by incentive condition..... 210

LIST OF SUPPLEMENTAL FILES

Supplemental Table 1: Methods Details for All Experiments.

Supplemental Table 2: Percentage of Participants Purchasing in Each Split Experiment and Selling Format.

ACKNOWLEDGEMENTS

This dissertation would not exist without the support, assistance, and encouragement of so many professors, peers, coauthors, friends, and family members. I am deeply grateful to all those who helped make this journey possible.

First and foremost, I would like to thank my dissertation and doctoral advisor, On Amir, who has had a tremendous impact on my work and life. I admire his seemingly boundless knowledge of everything from countless academic literatures to the intricacies of the Marvel universe. He has taught me the values of curiosity, passion, and enthusiasm in research, and the infectious joy and sheer excitement of discovering something new. His expertise and patience with me have been invaluable. He has taught me how to dream up interesting questions, design creative experiments, and make collaboration engaging and fun. He worked tirelessly to promote me and my work, and without his generous advocacy I would not be where I am now, about to step into a role as a professor. On is my role model not only in academia, but also in balancing work and play to craft a wonderful life. I have been truly honored to be his student and am eternally grateful for his guidance and support.

Ayelet Gneezy and Wendy Liu have also been fantastic coauthors and advisors. Ayelet is one of the most protective and caring professors I know. I could not see myself here today without her taking the time to check in with me, helping me navigate the personal and academic challenges of grad school, helping me discover my research interests, inviting me to collaborate on a fun and impactful project, and believing in me. I am forever grateful that she took the time to know me not only as a researcher but as a

person. Wendy has similarly been a beacon of support and encouragement—her positivity and resilience have taught me that even when research is tough, there is always a bright side and a new opportunity. Her talent for coming up with clever ideas, and for framing and positioning them, has also helped make me a better researcher.

The other members of my committee also deserve particular thanks. I am lucky that my advisor has good taste in friends, which allowed me to benefit from Jonathan Levav's brilliance and kindness. He patiently provided feedback on my papers and his suggestions greatly improved my work. I am also grateful for his advice, particularly about navigating the job market and not sweating the small stuff. I also thank Piotr Winkielman, one of my first professors at UCSD. I still use the knowledge I learned in his social psychology class, which was enriched with his entertaining and poignant anecdotes and discussions. I greatly appreciate the perspective he brings. All of the members of my committee have contributed to and helped improve this dissertation in various ways.

I am also grateful for the other current and former Rady faculty who encouraged me and helped me throughout grad school, especially while preparing for the job market: Uri Gneezy, Karsten Hansen, Uma Karmarkar, Michael Meyer, Kanishka Misra, Vincent Nijs, Chris Oveis, Yuval Rottenstreich, Sally Sadoff, Robert Sanders, David Schkade, Marta Serra-Garcia, Pam Smith, Charlie Sprenger, Ken Wilbur, and of course Nora Williams. Their feedback made my work more interdisciplinary, thoughtful, and interesting. I am also thankful for Pam's dedication to the PhD students in hosting the SOW/SW/W lab—a venue for sharing projects, discussing ideas, practicing presentations, and fostering community. I also thank the staff members who made Rady a great place to

be, including Patricia, Chau, Erwin and his facilities staff, Marisol Nierva-Magnano, Hillary Flocke, and others. I also benefited greatly from the work done by Dean Sullivan, who helped shape Rady into the amazing institution it is today.

I would also like to thank my colleagues and friends, at UCSD and around the world. I am so thankful that Allie Lieberman chose to join UCSD, becoming my classmate, officemate, conference roommate, coauthor, and friend. We learned together, laughed together, stressed together, brainstormed together, wrote together (word-by-word), and shared this whole wild experience together. I am also thankful for the kindness and thoughtfulness of my other Rady peers, including Min Zhang, Yidan Yin, Giulia Maimone, Jean Zhang, Ariel Friedman, Gal Smitizsky, Matt Lupoli, Silvia Saccardo, and especially Liz Keenan and Coby Morvinski who trailblazed this path. A special thanks is owed to my friend Evan Weingarten who has fascinated me with his encyclopedic brain. I am also thankful for the fantastic lab managers who helped make my research possible, dutifully executing even my most complicated, wacky studies: Steph Schwarz, Yael Horwitz, Jyoti Jhita, and Tiss Doci.

I am supremely grateful for my academic peers who braved the job market with me and helped keep me sane, Melanie Brucks and Rachel Gershon. They have been wonderful friends and are sure to be wonderful future coauthors. I am also thankful for my "econs": Pablo, Jason (the nucleus), Seung, Leland, Grant, and the Alexs. Thank you for making my time in San Diego so fun and for enriching my understanding of people and the world through the lens and lingo of economics. Many thanks also to my amazing crag crew—Kristina, Edwin, Xavi, Brooke, and Steve—thank you for helping push me to become a

better and stronger climber, the perfect hobby to keep my mind off work when I most needed it. Thank you also to my amazing best friends from back home who cheered me on every step of the way and have been an endless source of support and love: Kristie, Gretchen, Natalie, Aaron, and Danny.

Finally, no words can do justice to how grateful I feel for my family, my partner, and my future family. My mom, dad, brother, and grandparents have been my biggest supporters in every step of my life, always encouraging me with beaming pride. Thank you for believing in me, supporting me, and loving me. Grandma and Grandpa Duke and Grandpa Mileham: I know you are smiling down and would be so proud and joyful to share in this. Steve, Anne, Shane, Lorren, Randy, Tony, and my other family members have also been a great source of support and encouragement. My soon-to-be new family, Laura, Jim, and Trevor, have been so supportive and giving of their time, advice, and love, and I am thrilled to join their family soon. And last but certainly not least, I thank Tyler Liberty, my fiancé. Thank you for sharing this life with me and supporting me through this wild process—for the countless conversations about research, brainstorming materials, helping me prepare presentations, tolerating my busyness—and for your love and all the sacrifices you have made to make this possible. I am so fortunate.

Chapter 1, in full, is a reprint of previously published material as it appears in the Journal of Consumer Research, 45 (6): 1254—1273, 2019, Duke, Kristen Elizabeth and On Amir. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Duke, Kristen Elizabeth and On Amir. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of previously published material as it appears in Organizational Behavior and Human Decision Processes, 151: 118—131, Lieberman, Alicea, Kristen Elizabeth Duke, and On Amir. The dissertation author was one of the primary investigators and authors of this paper.

VITA

2013	Bachelor of Arts, Economics and Psychology, The College of New Jersey
2013-2019	Teaching Assistant, Rady School of Management, University of California, San Diego
2019	Doctor of Philosophy, Management, University of California, San Diego

PUBLICATIONS

Duke, Kristen E. and On Amir (2019), "Guilt Dynamics: Consequences of Temporally Separating Decisions and Actions," Journal of Consumer Research, 45 (6), 1254—73.

Lieberman, Alicea, Kristen E. Duke, and On Amir (2019), "How Incentive Framing Can Harness the Power of Social Norms," Organizational Behavior and Human Decision Processes, 151, 118—31.

Duke, Kristen E., Kelly Goldsmith, and On Amir (2018), "Is the Preference for Certainty Always So Certain?," Journal of the Association for Consumer Research, 3 (1), 63—80.

Ward, Adrian, Kristen E. Duke, Ayelet Gneezy, and Maarten Bos (2017), "Brain Drain: The Mere Presence of One's Own Smartphone Reduces Available Cognitive Capacity," Journal of the Association for Consumer Research, 2 (2), 140—154.

Crawford, Jarret T., Sophie A. Kay, and Kristen E. Duke (2015), "Speaking out of Both Sides of Their Mouths: Biased Political Judgments Within (and Between) Individuals," Social Psychological and Personality Science, 6 (4), 422—30.

ABSTRACT OF THE DISSERTATION

How the Structure of Choices Influences Consumer Decisions and Experiences

by

Kristen Elizabeth Duke

Doctor of Philosophy in Management

University of California San Diego, 2019

Professor On Amir, Chair

This dissertation contains three papers that together demonstrate how the structure of a choice process and of the information consumers encounter can influence what they think, feel, and do. Chapter 1 investigates how the structure of a choice process—and in particular, the timing of a decision to act relative to the timing of the action—influences how individuals feel at the point of acting. We demonstrate that the emotional experience

of guilt is composed of two dissociable sources: one tied to the decision to act, and one tied to the behavior itself. With this unpacking, we find that introducing a temporal gap between one's decision to act and one's behavior allows decision guilt to decay, which reduces the guilt at the point of acting and carries important behavioral consequences.

Chapter 2 probes the structure of a purchasing interaction. We compare two ways that the purchase and quantity decisions in a purchase process can be organized. In one, customers make the purchase and quantity decisions separately; in the other, customers simultaneously indicate whether and how much to buy. We find that the simultaneous consideration of both choices changes the psychology of how people decide, and ultimately increases their willingness to purchase. In 27 lab experiments and a large field experiment, we find that this "quantity integration" can lead to substantial lift.

Finally, chapter 3 reveals how the structure of one aspect of a customer interaction, a small monetary incentive, can influence what people believe and accordingly what they do. We show that customers who encounter a small surcharge for failing to perform some behavior infer that this behavior is more socially normative than those who encounter a small discount for performing it. This inference carries key consequences, changing how people feel and what they decide to do in the future. Collectively, these three chapters demonstrate the critical importance of the structuring of choices and information for consumers' beliefs, emotional responses, decisions, and behaviors.

Chapter 1.

GUILT DYNAMICS: CONSEQUENCES OF TEMPORALLY SEPARATING DECISIONS AND ACTIONS

Kristen Elizabeth Duke, On Amir

Rady School of Management, University of California, San Diego, La Jolla, CA, 92093, USA

ABSTRACT

The current research demonstrates that temporally separating a consumer's initial decision to perform a guilt-inducing action from its actual enactment reduces the guilt felt while acting. This hypothesis follows from the development of a dynamic model that unpacks guilt into two distinct components. Initially, one experiences decision guilt accompanying the decision to act or the realization that one will act; subsequently, one experiences action guilt while engaging in the guilt-inducing behavior. Four experiments and two pilot studies reveal that introducing a temporal "decision-enactment gap" enables decision guilt to decay in this interim period, which lowers the overall guilt experienced upon acting. In line with the self-regulative function of guilt, decision-enactment gaps also increase indulgent consumption and decrease post-behavior atonement. This decoupling process can thus alleviate guilt that might otherwise detract from experiences, but may come at a cost to self-control efforts. The authors discuss the theoretical and practical implications of these findings.

INTRODUCTION

Consumers perform many behaviors that can induce guilt. This unpleasant emotion may arise when splurging on an expensive product, skipping a workout, breaking one's diet, declining a request to donate, discarding a recyclable item, purchasing from a disreputable business, and beyond. Prior research demonstrates that guilt peaks after one engages in such actions, and dissipates gradually thereafter (Macht and Dettmer 2006; Ramanathan and Williams 2007). For example, a dieting individual may feel a pang of guilt after consuming a decadent dessert, but this feeling tends to decay over time. In the current research, we advance the novel hypothesis that guilt can decay not only after, but also prior to acting.

To arrive at this conclusion, we develop a dynamic model in which the overall guilt a consumer experiences during a behavior is a composite of two sources: decision guilt and action guilt. We define decision guilt as that associated with the decision or realization that one will commit a guilt-inducing act, and action guilt as that associated with engagement in the actual action. We posit that decision guilt decays over time. Therefore, when there is temporal separation between one's initial decision or realization and the eventual behavior, decision guilt can decay in the interim period. As a result, temporal "decision-enactment gaps" reduce the remaining amount of decision guilt—and accordingly, the overall level of guilt—consumers feel upon acting.

In addition to demonstrating the basic effect of decision-enactment gaps on how guilty consumers feel, the current research also identifies two important behavioral consequences. First, drawing on the idea that guilt prompts individuals to cease their guilt-

inducing behavior (e.g., Monteith et al. 2002), and that decision-enactment gaps reduce guilt during consumption, we find that decision-enactment gaps increase indulgent consumption. Second, decision-enactment gaps also reduce atonement behaviors after the action has been performed (e.g., Tangney et al. 1996). Thus, not only can a decision-enactment gap reduce how guilty consumers feel, it can also increase how much they treat themselves and decrease how much they make up for their indulgence.

In the following sections, we review prior research relevant to the present framework and develop our hypotheses. Thereafter, we test these ideas in two pilot studies and four experiments. Finally, we conclude by addressing the theoretical and managerial implications of these findings.

GUILT AND THE ROLE OF PERSONAL RESPONSBILITY

Guilt is an unpleasant emotion experienced in response to a negative evaluation of one's current state (Lascu 1991; Smith and Ellsworth 1985). It arises when an individual perceives that he/she has fallen short of internalized standards of appropriate, desired conduct (Baumeister, Stillwell, and Heatherton 1994). Two important features differentiate guilt from other negative emotions. First, guilt is linked to specific actions that involve actual or perceived harm—either harm to others, as when lying to a friend (Baumeister et al. 1994), or harm to the self, as when breaking a diet, buying an expensive product, or not studying enough (Dahl, Honea, and Manchanda 2003; Mukhopadhyay and Johar 2009; Tangney 1992). Unlike similar negative emotions such as shame, guilt is associated with a

sharpened focus on the consequences of a specific behavior (Lewis 1971; Tracy and Robbins 2004), rather than a broader, more global evaluation of the self.

Second, guilt arises from attributions of personal responsibility and self-blame (e.g., Smith and Ellsworth 1985; Tangney et al. 1996). As such, guilt is a "self-caused" emotion resulting from a cognitive appraisal regarding one's own actions, and is similar in this respect to pride or shame (Roseman 1984). In fact, the link between guilt and personal responsibility is so strong that merely activating the emotion of guilt can enhance feelings of personal control over unrelated, uncontrollable events (Kouchaki, Oveis, and Gino 2014). Guilt is also different in this manner from other-caused emotions, such as anger, which arise following attributions about the actions of other people (Neumann 2000), and from externally-caused emotions, such as fear (Roseman 1984), that arise when a focal event is caused by circumstances outside one's control (though certainly, fear may also follow the actions of other people). In sum, what distinguishes guilt from other emotions is that it arises when individuals focus on specific behaviors for which they feel personally responsible, and which are expected to harm oneself or someone else.

To illustrate this distinction, imagine that one friend in a group chooses a restaurant at the other end of town, only to realize upon arriving that it is closed for renovations. All members of the group might feel anger, sadness, or other emotions directed at the restaurant, the situation, or the friend who made the selection. However, it is only the friend who chose that restaurant—that is, the person with personal responsibility for the harm-inducing action—who might feel guilty about having chosen it.

We use the perspective that personal agency is central in rousing guilt to inform the idea that initial realizations that one will act may also be guilt-inducing, even in advance of the action. Consider an individual who consumes an unhealthy slice of cheesecake, despite being on a diet. She may feel guilty not only because of her actual consumption (e.g., the number of calories consumed and their impact on her waistline), but also for the very fact of having chosen to eat it. Said otherwise, she may feel guilty about making the initial unhealthy decision to break her diet, even before she takes the first bite. This guilt is tied to the knowledge that one will act in a guilt-inducing manner and can accordingly be experienced before acting. As mentioned previously, we term this construct decision guilt.

One should note that decision guilt is distinct from anticipatory guilt. The former is, as we propose, an affective experience that arises after a choice has been made to act in a guilt-inducing manner. The latter is a cognitive awareness that one will feel guilty in the future, which arises during the decision-making process and can influence how consumers choose to act (Baumeister et al. 2007; Lindsey 2005; Peloza, White, and Shang 2013). For example, when deciding whether to buy an indulgent dessert, a consumer may consider how she might feel after eating it, forecasting her post-consumption state (Rick, Cryder, and Loewenstein 2008; see also Loewenstein et al. 2001 and Pham 2004 for broader discussions of anticipated emotions). Note that a consumer may also experience anticipated guilt after a decision has been made, if she considers how she will feel upon acting. Yet, this metacognitive knowledge that guilt will arise in the future is still distinct from decision

guilt, the visceral emotion experienced right when a decision has been made. Pilot Study 1 demonstrates that decision guilt is indeed distinct from the anticipated guilt that can arise either prior to a decision (anticipatory guilt) or following a decision (at the same point in time as decision guilt).

In addition to decision guilt, a second component of consumer guilt is what we term action guilt. Action guilt arises once an individual engages in a guilt-inducing action—for example, when she eats an unhealthy food, acts rudely to another person, throws away a gift, or splurges on a frivolous purchase (Dahl et al. 2003). Thus, it aligns with the traditional concept of guilt in consumer research, arising when consumers behave in a manner perceived to cause harm (e.g., Baumeister et al. 1994; Dahl et al. 2003).

To summarize, we propose that a consumer's overall experience of guilt associated with a given behavior can be unpacked into two components: decision guilt that arises in association with the realization that one will act, and action guilt that arises in association with the actual action. (Note, however, that the current research does not investigate the guilt that may arise after a period of rumination following the behavior, e.g., Rachman 1993.) This decomposition of guilt into two components prompts a set of novel predictions as to the dynamic nature of guilt and its corresponding impact on consumers' experiences and behaviors.

THE TEMPORAL DYNAMICS OF GUILT

Emotions are dynamic experiences that can decay over time (Ekman 1984; Gilovich, Medvec, and Kahneman 1998; Jager et al. 2014; Kahneman 1995), and guilt is no exception (e.g., Kivetz and Keinan 2006). Prior work investigating the temporal profile of guilt suggests that it peaks shortly after a behavior and decays over time thereafter (Macht and Dettmer 2006; Ramanathan and Williams 2007). Our theoretical framework predicts that decision guilt, the specific component of overall guilt that is linked to initial realizations that one will act, will also decay over time (and importantly, in the interim period between decision and action).

A variety of psychological and physiological processes likely contribute to this decay (Gilbert et al. 2004; Taylor 1991). Consumers may consciously rationalize their decisions (Dahl et al. 2003), engaging in explicit emotion regulation to re-evaluate the event's underlying meaning and reduce its emotional impact (e.g., Gross 1998). This rationalization process may also operate passively and unconsciously (Lieberman et al. 2001; Gyurak, Gross, and Etkin 2011). Emotional crowding-out can also play a role: a consumer who decides to buy an expensive purse and feels guilty about it may proceed to answer work-related emails, go rock climbing, listen to a podcast, and so on. The relative salience of guilt may fade, pushed aside by the more salient and vivid emotions elicited by such events (Izard 1977). Unconscious, biological adaptation and habituation processes may also work to reduce guilt over time. That is, the moderate physiological arousal response associated with guilt (Baumeister et al. 1994; Smith and Ellsworth 1985) may organically return to homeostatic baseline (e.g., Sandvik, Diener, and Larsen 1985), attenuating this emotion. Furthermore, the salience and vividness of the memory of one's

personal responsibility for the focal decision (i.e., the appraisal associated with guilt) may decay over time, and accordingly the guilt-eliciting cue may decay as well. Note that which mechanism operates most strongly is likely dependent on the context, the length of the decision-enactment gap, and the intensity of guilt. Regardless of which process is most responsible, we propose and find that decision guilt decays over time (Pilot Study 2). Accordingly, when time elapses between initial realizations and actual behaviors, these decision-enactment gaps should allow decision guilt to decay prior to acting. As a result, decision-enactment gaps should also reduce consumers' overall guilt when acting.

GUILT DYNAMICS FRAMEWORK

Figures 1.1 and 1.2 present our overall theoretical framework, with curves depicting predicted levels of guilt over time. While there are several possible shapes for such curves, for illustration purposes, they are drawn with simplifying assumptions. First, these figures depict a case in which decision guilt and action guilt reach approximately equal peak levels, although our central predictions would also hold when decision guilt exceeds action guilt, or when action guilt exceeds decision guilt.[1]

Second, we map the decay of guilt as a smooth, continuous process (analogous to that of short-term memory, e.g., Peterson and Peterson 1959). Note that our results would

[1] Consider that the guilt a consumer experiences while acting is "overall guilt," a composite of (decayed) decision guilt and action guilt. Even when the initial decision guilt is low, it should still decay. As a result, decision guilt after a decision-enactment gap should be lower than decision guilt absent such a gap, with the former producing lower overall guilt upon acting. That said, although mathematically our predictions would hold even when decision guilt is low and action guilt is high, the decay of decision guilt would not make a large difference, proportionally, on overall guilt in this situation. Such a difference may not be large enough to measure using traditional tools (e.g., self-report).

also hold if the decay of guilt follows a different function (e.g., a step-wise decay); the only requirement for our predictions is a decrease in the level of guilt felt over time. Finally, these figures depict decision guilt and action guilt as additive components of overall guilt. However, our results would still hold with sub-additivity between these components, or with other forms of interaction. The only requirement for our analysis is that when decision guilt and action guilt are experienced in close temporal proximity, the resultant overall guilt experienced is higher than the overall guilt experienced when decision guilt has had time to decay before the action begins.

The graphs in Figure 1.1 depicts two key tenets of our theory. First, the global decay of guilt over time is indicated by the downward trend in guilt as time progresses. Second, our representation of overall guilt as a composite of decision guilt and action guilt is depicted such that the black curve in each graph is a composite of the gray curves. Importantly, the left graph depicts a scenario in which the realization that one will act and the enactment of that action occur in close temporal proximity. Thus, in this scenario, decision guilt and action guilt are initiated at similar points in time (i.e., there is no decision-enactment gap). In contrast, the right graph depicts a scenario in which time elapses between the initial realization and the action (i.e., there is a decision-enactment gap). According to our theoretical framework, such a situation allows decision guilt to decay prior to the action. As a result, the overall level of guilt a consumer experiences while acting should be lower than it was in the left graph, where there was no decision-enactment gap.

Next, Figure 1.2 directly compares the overall guilt predicted to arise with and without a decision-enactment gap. As shown, our framework predicts that individuals should experience less guilt when there is a gap. Note that we investigate the level of guilt, rather than consumers' cumulative experience of guilt (the total area under the overall guilt curve), for which we make no prediction.

In summary, we propose that consumers experience a pang of guilt upon the initial realization that they will act (decision guilt). This emotional experience decays over time. Thus, when time passes between a decision and the action (i.e., a decision-enactment gap), this decision guilt can decay and contribute less emotional distress to consumers' experiences while carrying out the action, so that they will experience a lower level of guilt.

In this analysis, then, we conceptualize a process in which consumers may "emotionally pre-pay" their guilt by reaching a decision to act and allowing time to pass before actually acting ("paying" their decision guilt in the interim period). This process bears resemblance to a component of intertemporal mental accounting (Prelec and Loewenstein 1998; Thaler 1985). Research in that domain has shown that consumers experience a pain of payment (e.g., when paying for an expensive vacation), and so they prefer to pre-pay financial costs prior to the experience to prevent this pain from undermining their enjoyment (Prelec and Loewenstein 1998). A key feature of this dynamic is that consumers gradually adapt to financial costs that they have paid with the passage of time (Gourville and Soman 1998). Just as consumers adapt to financial payments, then, we posit that they also adapt to emotional costs, and that decision-

enactment gaps enable "pre-payment" of these emotional costs so that consumption experiences can be enjoyed with less guilt.[2]

THE BEHAVIORAL CONSEQUENCES OF GUILT

The current framework has implications not only for how consumers feel, but also for how they act. Guilt affects behavior in two phases: interruption and then reparation (Amodio et al. 2007). Initially, the experience of guilt encourages consumers to interrupt and cease their ongoing guilt-inducing behavior (Monteith et al. 2002). That is, experiencing guilt tends to reduce individuals' approach motivation (Amodio et al. 2007) and inhibit negative action, such as by reducing consumption of unhealthy food (Giner-Sorolla 2001).

Guilt also performs a critical function after the behavior has already occurred: promoting atonement. Particularly, guilt motivates engagement in reparative actions with the goal of amending the harm imposed (Lindsay-Hartz, De Rivera, and Mascolo 1995; Miceli and Castelfranchi 1998; Tangney et al. 1996). For example, a consumer who feels guilty for eating an indulgent meal may decide to hit the gym in an attempt to repair the damage to her health, and a consumer who feels guilty for letting down a friend might purchase a gift in hopes of repairing the relationship. Thus, after the action has been

[2] However, note the key difference between separating two distinct actions (e.g., payment and consumption), as in the mental accounting literature, and separating a decision from its linked action, as in the current research.

performed, guilt promotes approach motivation oriented toward repairing this harm (Amodio et al. 2007).

Drawing on these empirical foundations, decision-enactment gaps should not only reduce guilt when acting, as discussed, but also (1) diminish the interruptive function of guilt, allowing consumers to behave more freely during the experience, and (2) diminish the atonement function of guilt, reducing consumers' willingness to engage in reparations after acting.

OVERVIEW OF THE EXPERIMENTS

Two pilot studies and four experiments test our theoretical framework. Pilot Study 1 provides evidence of decision guilt, establishing that consumers indeed feel guilty upon learning that they will engage in a harmful behavior—prior to the actual action. Further, it demonstrates that decision guilt is a unique emotional experience, distinct from anticipated and anticipatory guilt. Pilot Study 2 then establishes that decision guilt decays over time (in the period prior to acting). The remaining experiments test for the consequences of decision-enactment gaps, and all follow the same general structure, which is depicted in Figure 1.3. Participants make a decision (or come to a realization), complete filler tasks, perform their chosen action, and complete various focal measures. For certain participants, the filler tasks arise in between the initial decision/realization and the action, thus producing a decision-enactment gap.

Experiment 1 find that decision-enactment gaps reduce consumers' guilt upon acting, but not other negative emotions such as shame and regret, thus providing discriminant validity. Next, Experiment 2 employs a repeated-measures design to link the results of Pilot Study 2 with those of Experiment 1. Specifically, it identifies a relationship between the decay of decision guilt during a decision-enactment gap and the resultant lower guilt upon acting.

Following, Experiment 3 draws on the behavior-interruptive function of guilt to find that decision-enactment gaps increase the amount of unhealthy food consumed. Finally, Experiment 4 demonstrates that, in line with the restorative function of guilt, decision-enactment gaps reduce consumers' atonement for their guilt-inducing behavior. In all experiments, we report all manipulations and all measures. For conciseness, we include information on how we determined sample sizes and details on data exclusions (if any) for all studies, as well as additional methodological details, in the Appendix.

PILOT STUDY 1: IDENTIFYING DECISION GUILT AND DISTINGUISHING IT FROM ANTICIPATED GUILT

Pilot Study 1 tests whether consumers can feel guilty upon mere realizations that they will act in a guilt-inducing manner. Participants learn that they will steal from another participant, and are led to believe that their actions played a part in this role determination. After realizing they will need to steal, but prior to the actual theft (i.e., before carrying out the guilt-inducing action), participants report how guilty they currently feel. This post-

realization value is compared to a baseline condition in which individuals report their guilt prior to the realization, to demonstrate that such a realization indeed induces (decision) guilt.

Further, this decision guilt is also compared to the guilt participants anticipate feeling upon stealing, measured at two points in time: either (a) before the realization (i.e., "anticipatory guilt" as defined in prior literature), or (b) after the realization, but before the theft (i.e., at the same point as decision guilt, but rather than capturing guilt felt in that moment, instead capturing the guilt participants expect to feel when they carry out the action). We propose that decision guilt is a distinct emotional experience and will accordingly significantly differ from anticipated guilt (a cognitive awareness of one's future emotions) measured at either point in time.

Method

Overview. Two hundred forty-two undergraduates (59% female, M_{age} = 21.2 years, SD_{age} = 2.2 years) participated for partial course credit. Participants were randomly assigned to one of four conditions (baseline, anticipatory guilt, decision guilt, or post-decision anticipated guilt), between-subjects. All participants completed the same study and responded to a set of measures assessing their emotional state, but the timing of this measurement varied across conditions.

Game Introduction. At the beginning of the experiment, baseline participants read, "Before we begin, we are interested in how you feel in this moment. Right now, to what

extent do you feel…guilty?" (1: Not at all to 9: Very much). Then, all participants learned that they would play a game and responded to an attention check (see Appendix). Thereafter, participants chose one of two options labeled "left" or "right," which would assign them to play as either a rich or a poor player for the duration of the game. Participants were unaware which option corresponded to which wealth level. (The Appendix presents the full text presented to participants for this and all remaining game situations and choices.)

Regardless of which option they had clicked, the next page read, "You chose the left [right] option. This option assigns you to the rich wealth level. This means you will play as a rich player. You will have an advantage over other lab participants playing as poor players." On the following page, participants read that poor players would complete a challenging task, but some would get lucky and receive an "inheritance" ($50). At the end of the week, one participant would be selected to receive any money won in this game. Rich players would either rob a poor person (and steal the inheritance money, if that player had one) or donate to a poor person: receive $25 themselves and give an additional $25 to a poor player. They learned they would "choose" their game strategy in the same manner that they had "chosen" to play as poor or rich.

At this point, anticipatory guilt participants reported how they expected to feel if they stole: "Please imagine that you chose to rob the poor. Now, imagine you have completed this role—that is, you selected a poor player to rob out of a list of SONA IDs [anonymous, unique participant identifiers]. At that moment, to what extent would you feel…guilty?" (1: Not at all to 9: Very much).

Role Determination. On the following page, all participants chose one of two options labeled “top” or “bottom,” which would ostensibly assign them to either rob or donate. Here again, participants were unaware which box corresponded to which action. Regardless of what they chose, the next page read, “You chose the top [bottom] option. This option is rob the poor. This means you will select a poor player to rob out of a list of SONA IDs.”

At this point, decision guilt participants reported how they currently felt: “Right now, to what extent do you feel…guilty?” (1: Not at all to 9: Very much). At the same point in time, post-decision anticipated guilt participants reported how they expected to feel after stealing: “Now, imagine you have completed the role of robbing the poor—that is, you selected a poor player to rob out of a list of SONA IDs. After stealing, at that moment, to what extent would you feel…guilty?” (1: Not at all to 9: Very much).

Stealing. Finally, all participants selected the poor player they would rob by choosing a SONA ID from a drop-down list and provided demographic information.

Results

Figure 1.4 displays average levels of guilt (experienced or anticipated) in each condition. Supporting the existence of decision guilt, individuals in the decision guilt condition felt significantly guiltier ($M = 4.62$; $SD = 2.83$) than did baseline participants ($M = 2.10$, $SD = 1.78$; $t(119) = 5.87$, $p < .001$).

Further supporting our predictions, decision guilt also significantly differed (M = 4.62, SD = 2.83) from both anticipatory guilt (measured prior to the realization; M = 6.23, SD = 2.42; $t(119) = 3.37$, $p = .001$) and post-decision anticipated guilt (measured after the realization; M = 6.37, SD = 3.04; $t(118) = 3.26$, $p = .001$). The two anticipated guilt measures did not differ from one another ($t(119) = .27$, $p = .784$).

Discussion

Pilot Study 1 demonstrates that individuals feel guilty upon realizing they will commit a guilt-inducing action—that is, they experience decision guilt. Further, decision guilt was different from the guilt individuals anticipated feeling upon carrying out the action, when measured either at the same point in time as decision guilt or in the earlier decision-making phase. These findings demonstrate that decision guilt is not a manifestation of forecasted action guilt; rather, it seems to be a distinct emotional experience. According to our theoretical framework, decision guilt should also decay over time. We test this possibility in Pilot Study 2.

PILOT STUDY 2: DEMONSTRATING THE DECAY OF DECISION GUILT

As in Pilot Study 1, participants learn that they will steal from another participant, and are led to believe that their actions played a part in this role determination. Prior to the theft (i.e., before carrying out the guilt-inducing behavior), participants report on their

emotional states, either immediately upon learning of their role or after time has elapsed. We predict that consumers will again report experiencing decision guilt. Importantly, we also predict that this emotion will decay over time. However, a decrease in reported guilt after a delay may reflect other unrelated dynamics. Participants could report lower emotions throughout the course of a study because of scale adaptation; similarly, the awareness of their role could increase all reported emotions, which would then decline back to baseline after a break. To reject these alternative accounts for the predicted pattern of guilt, participants also report a set of additional emotions; we show that these emotions do not reveal the same temporal pattern.

Method

Overview. One hundred ninety-nine Amazon Mechanical Turk workers (51% female, M_{age} = 36.9 years, SD_{age} = 11.8 years) participated for a small monetary payment. Participants were randomly assigned to one of two conditions (no-delay or delay), between-subjects. No-delay participants completed filler tasks, learned about the game (similar to Pilot Study 1, and described next) and flipped a coin to determine their role, and then reported their emotional state. Delay participants completed the same activities, but instead completed the filler tasks after the coin flip—before reporting their emotions. Thus, during the filler tasks, there was time for their emotional response to decay before it was measured. Finally, all participants stole the payoffs of another participant, per the game rules. We describe each of these actions next.

Game Role Determination and Filler Tasks. Participants learned that they would play a game and then responded to the same attention check as in Pilot Study 1 (except "lab participants" was replaced with "MTurk workers"). Next, they learned they would flip a coin to determine their role in the game. They would play the game either as a knight and "fight honorably" by solving puzzles, where the knight who solved more (vs. fewer) puzzles would earn a 50-cent (vs. 5-cent) bonus payment, or instead could play as a thief. Thieves would "select a knight and take all the money that knight has worked for, leaving that knight with $0." Participants then practiced flipping a (fair) virtual coin. On the following screen, participants flipped a second virtual coin that determined their role. Unbeknownst to them, this latter coin was not fair/random; all participants were assigned to the role of thief, and would thus need to steal the bonus earnings of another participant.

As mentioned previously, participants also completed approximately five minutes of unrelated filler tasks (details in the Appendix). For no-delay participants, these filler tasks were completed before learning about the game. For delay participants, these filler tasks were completed after flipping a coin to determine their role of the thief; thus, there was a delay between the role revelation and when they reported their emotions.

Reporting Emotions. Next, in randomized order, participants indicated how ashamed, regretful, guilty, sad, fearful, and happy they felt "right now" (1: Not at all to 9: Extremely). These emotions were selected to match certain characteristics of guilt. Shame and regret are both self-caused emotions that are negative in valence, like guilt (e.g., Roseman 1984). Sadness and fear were included as control emotions that are negative in valence, but not self-caused. Finally, happiness was included as a control emotion that is

positive in valence. For no-delay participants, as mentioned, this reporting occurred immediately after learning that they would play as a thief. Thus, any emotions measured at this point reflect the emotions experienced upon the initial realization that they would engage in a harmful behavior (stealing), before the theft. For delay participants, this reporting occurred after the filler tasks that followed their role revelation. This delay should allow emotions to dissipate before being reported.

Stealing. Next, all participants were directed to steal the bonus earnings of one participant playing as a knight. They chose which participant they would steal from by selecting a participant ID from a drop-down list (these IDs were fictitious, but followed the structure of typical MTurk worker IDs). On the following screen, participants learned that they stole a 50-cent bonus from this selected MTurk worker. Finally, participants answered, "How responsible do you feel about MTurk worker [ID] losing his/her earnings?" (1: Not responsible at all to 9: Completely responsible) and provided demographic information.

Results

Guilt. Supporting the idea that consumers experienced decision guilt that decayed over time, no-delay participants reported feeling significantly higher guilt ($M = 3.16$, $SD = 2.65$) than did delay participants ($M = 1.99$, $SD = 1.77$; $t(197) = 3.67$, $p < .001$).

Other Emotions. We tested whether other emotions showed the same pattern as guilt. Like guilt, albeit with a smaller decline, participants reported higher shame in the no-

delay condition (M = 2.84, SD = 2.47) than in the delay condition (M = 2.09, SD = 2.01; t(197) = 2.35, p = .020). However, this reduction in shame following a delay is not robust across the remaining studies. Furthermore, participants were similarly regretful (no-delay: M = 3.21, SD = 2.67; delay: M = 2.79, SD = 2.47; t(197) = 1.16, p = .248), sad (no-delay: M = 2.94, SD = 2.26; delay: M = 2.71, SD = 2.42; t(197) = .69, p = .490), fearful (no-delay: M = 2.47, SD = 2.06; delay: M = 2.16, SD = 2.03; t(197) = 1.09, p = .279), and happy (no-delay: M = 5.38, SD = 2.46; delay: M = 5.19, SD = 2.33; t(197) = .57, p = .569) in both conditions, suggesting no significant reduction in these emotions following a delay.

Discussion

Pilot Study 2 reveals that decision guilt decays with the passage of time. But what is the significance of decision guilt and its decay in the period prior to acting? Does it contribute, as we have posited, to a reduction in the overall level of guilt consumers experience upon acting, or is it swamped by action guilt, leaving no measurable influence? The next experiment addresses these questions by testing whether decision-enactment gaps reduce overall guilt upon acting.

EXPERIMENT 1: DECISION-ENACTMENT GAPS REDUCE GUILT UPON ACTING

As in the pilot studies, Experiment 1 participants are led to steal the earnings of another participant either immediately after learning that they have earned this role, or after

a delay. However, unlike the pilot studies, this experiment examines participants' overall guilt immediately after engaging in the guilt-inducing action of stealing. We predict that decision-enactment gaps will reduce participants' overall guilt upon acting. In addition to testing for this effect, Experiment 1 also seeks discriminant validity by testing whether decision-enactment gaps reduce other emotions beyond guilt.

Method

Overview. One hundred twenty-one undergraduates (33% female; M_{age} = 21.5 years, SD_{age} = 1.9 years) participated for partial course credit. Participants were randomly assigned to one of two conditions (unseparated or separated) between-subjects. All participants completed four activities in this experiment. Separated participants learned about the game (similar to Pilot Study 2, and described next) and flipped a coin to determine their role, completed filler tasks, stole money from another participant, and finally reported their emotions. Thus, there was a decision-enactment gap between the realization (the coin flip) and the behavior (stealing), immediately after which emotional states were assessed. Unseparated participants completed the same activities, but instead completed the filler tasks before the coin flip; thus, there was no temporal separation between the initial realization that they will steal and the actual theft.

Game Role Determination and Filler Tasks. Participants were taken one-at-a-time to a side room, where they learned they would participate in a study that involved creating audio recordings. Consenting participants remained in the room and began the study;

participants who did not consent were dismissed. Separated participants learned they would play a game alongside other participants and completed the same attention check as Pilot Study 1. At the end of the week, one participant would be randomly selected to receive his/her full game earnings as cash. Participants could either play as a knight and "fight honorably" by solving puzzles, with the knight who solved more puzzles earning $50 and the other earning $25, or instead could play as a thief. Thieves would "select a knight and take all the money that knight has worked for, leaving that knight with $0." Participants then practiced flipping a (fair) virtual coin.

Next, to enhance the social realism of the game, ensure comprehension, and increase engagement, participants made a verbal recording explaining the setup of the game in front of a research assistant. They read, "You should imagine you are explaining to another person what just happened and what will happen in the game. You should describe: *how many players are in the game, *the roles of the players, *the potential actions and consequences in the game, *etc." On the following screen, participants flipped a second virtual coin that determined their role. Unbeknownst to them, this latter coin was not fair/random; all participants were assigned to the role of thief, and would thus steal the earnings of another participant. Thereafter, participants made a second verbal recording. The instructions were, "You should imagine you are explaining to another person what just happened in the game and how you feel about it. Describe what happened and how you feel in your own words." As with the previous recording, this was designed to reinforce the role determination and heighten participants' engagement in the game.

After this role determination, separated participants completed unrelated filler tasks (approximately 5-8 minutes long; details in the Appendix). As mentioned previously, unseparated participants completed this same set of activities, but in an alternate order, completing the filler tasks before learning about the game and determining their role.

Stealing. Next, all participants were directed to steal the earnings of one participant playing as a knight. They selected a participant ID from a drop-down list (these IDs were fictitious, but followed the structure of typical IDs in the lab). On the following screen, they learned that they stole $50 in potential winnings from this selected participant. Thus, if they were chosen at the end of the week, they would receive $50. However, if the knight they stole from was chosen, the knight would learn that he/she won, but would receive nothing. Immediately thereafter, all participants made a third recording with the same instructions as the previous one: to explain what happened and how they felt.

Reporting Emotions. On the next page, participants reported the emotions they were currently feeling (the same emotions reported in Pilot Study 2). In randomized order, they indicated how sad, ashamed, regretful, guilty, fearful, and happy they felt "right now" (1: Not at all to 9: Extremely). Finally, participants answered, "How responsible do you feel about participant [ID] losing his/her earnings?" (1: Not at all to 9: Completely responsible) and provided demographic information.

Results

Guilt. Figure 1.5 displays all emotions by condition. In line with our predictions, participants felt significantly guiltier in the unseparated condition (M = 3.21, SD = 2.25) than in the separated condition (M = 2.40, SD = 1.95; t(119) = 2.11, p = .037). This result reveals that the decision-enactment gap indeed reduced participants' guilt upon acting.

Other Emotions. Although the decision-enactment gap reduced guilt, it did not reduce any of the other emotions measured. Participants did not report significantly different levels of shame in the unseparated (M = 2.36, SD = 1.56) and the separated conditions (M = 2.06, SD = 1.84; t(119) = .96, p = .337). Participants were also similarly regretful in both conditions (unseparated: M = 2.07, SD = 1.80; separated: M = 2.25, SD = 1.63; t(119) = -.59, p = .553), a pattern repeated with sadness (unseparated: M = 2.26, SD = 1.87; separated: M = 2.21, SD = 1.63; t(119) = .16, p = .870), fear (unseparated: M = 1.52, SD = .98; separated: M = 1.60, SD = 1.26; t(119) = -.42, p = .678), and happiness (unseparated: M = 5.02, SD = 2.13; separated: M = 5.16, SD = 2.26; t(119) = -.35, p = .724). These results suggest that the effect of a decision-enactment gap in this experiment was isolated to guilt.

Discussion

Experiment 1 reveals that decision-enactment gaps reduce how guilty individuals feel upon engaging in a guilt-inducing behavior. Together, Pilot Study 2 and Experiment 1 demonstrate decision guilt and its decay, and reveal the effect of decision-enactment gaps

on overall guilt upon acting. However, our theoretical framework supports a direct link between these two findings. We test this notion in Experiment 2.

EXPERIMENT 2: LINKING THE GUILT-REDUCING POWER OF DECISION-ENACTMENT GAPS TO THE DECAY OF DECISION GUILT

Experiment 2 employs a similar experimental setup as Experiment 1, but with a repeated-measures design. It tests for a link between the decay of decision guilt and the ability of decision-enactment gaps to reduce overall guilt upon acting.

Method

Overview. One hundred fifty undergraduates (60% female, M_{age} = 21.2 years, SD_{age} = 2.7 years) participated for partial course credit. Participants were randomly assigned to one of two conditions in a mixed design (between-subjects: separated or unseparated; repeated measure: feelings of guilt). All participants completed eight activities in this experiment, four of which were reporting their current level of guilt. All participants began by reporting their baseline levels of guilt. Thereafter, separated participants learned about the game (similar to that of Experiment 1, with slight variations described next), selected an option to determine their role, reported their decision guilt, completed filler tasks, reported their post-filler-task guilt, stole money from another participant, and finally reported their post-theft feelings of guilt. Thus, there was a decision-enactment gap

between the realization and the behavior (stealing). Unseparated participants completed the same activities, but instead completed the filler tasks before the role determination; thus, there was no temporal separation between the realization and the theft.

Details. At the beginning of the experiment, all participants read, "Before we begin, we are interested in how you feel in this moment. Right now, to what extent do you feel…guilty?" (1: Not at all to 9: Very much). The rest of the game followed the setup of Experiment 1 (filler task details are in the Appendix), with a few exceptions. Participants completed the study in a main computer lab with no audio recordings (participants only self-reported guilt). Rather than a coin flip, the role determination mirrored that of Pilot Study 1, where the choices of "left" or "right" and "top" or "bottom" determined participants' roles in the game. (Again, all participants ended up as thieves who must steal from another participant.) Participants also reported their guilt at three additional times throughout the experiment, in the same manner as baseline guilt.

Results

The Decay of Decision Guilt. Participants again experienced decision guilt, and it decayed over time. Specifically, participants reported significantly higher guilt after the realization that they would steal ($M = 3.98$, $SD = 2.77$) than they did at baseline ($M = 1.93$, $SD = 1.78$; paired $t(149) = 8.44$, $p < .001$). Furthermore, participants in the separated condition (i.e., those for whom the filler tasks followed the realization) reported significantly higher guilt immediately after the realization (i.e., decision guilt, $M = 4.21$,

SD = 2.77) than they did after the filler tasks that followed the realization (i.e., post-filler guilt, M = 2.49, SD = 1.91; paired $t(74) = 6.43$, $p < .001$; see Figure 1.6, left panel). This suggests that in the condition in which decision guilt could decay, it did. However, it did not decay all the way back to baseline (M = 1.76, SD = 1.45; paired $t(74) = 3.45$, $p < .001$), suggesting that on average in this experiment, a residual level of decision guilt remained. Consistent with this idea, this measure of post-filler guilt in the separated condition (M = 2.49, SD = 1.91) was marginally higher than the post-filler guilt reported in the unseparated condition (i.e., after the same filler tasks but absent a decision; M = 1.96, SD = 1.90; $t(148) = 1.72$, $p = .09$).

Decision-Enactment Gaps and Post-Action Guilt. We predicted that separated participants would report lower guilt upon stealing than unseparated participants would. To test this in accordance with the repeated-measures design, we first ensured that participants reported equal guilt at baseline. Indeed, participants reported similar levels of baseline guilt prior to the experimental manipulation (separated: M = 1.76, SD = 1.45 vs. unseparated: M = 2.09, SD = 2.05; $t(148) = 1.15$, $p = .253$). We also expected and observed no difference in the level of decision guilt participants felt after learning they would steal from a peer (separated: M = 4.21, SD = 2.77 vs. unseparated: M = 3.75, SD = 2.77; $t(148) = 1.03$, $p = .303$). Thus, to account for differences between participants attributable to their individual sensitivity to the guilt-inducing scenario (in line with this repeated-measures design), we control for these measures when comparing participants' post-action guilt. As one might expect, a regression predicting post-action guilt from decision guilt, baseline guilt, and condition (separated vs. unseparated) finds a significant positive effect of decision guilt (β

= .94, t(146) = 27.39, p < .001). There was no main effect of baseline guilt (β = -.03, t(146) = -.66, p = .513; excluding baseline guilt as a predictor in these analyses does not affect the results). Importantly, there was also a main effect of condition: participants reported significantly higher post-action guilt in the no-delay condition (least-squares M = 4.36, SE = .13) than in the delay condition (least-squares M = 3.86, SE = .13; β = .50, t(146) = 2.70, p = .008; see Figure 1.6, right panel). Thus, introducing a decision-enactment gap again reduced the level of guilt individuals felt immediately after performing the action.

The Link Between Decision Guilt Decay and the Effect of a Decision-Enactment Gap. We predicted that (a) decision-enactment gaps should reduce the level of guilt individuals feel upon acting because (b) such gaps allow decision guilt to decay. The results already presented support predictions (a) and (b) separately, but these findings should also be related to one another. Specifically, among separated participants, the level of post-filler guilt should reflect the residual level of decision guilt felt. Accordingly, it should directly correspond to the overall guilt participants feel upon acting: the lower the (residual) level of decision guilt, the lower overall guilt participants should feel upon acting (because overall guilt is a composite of action guilt and [decayed] decision guilt). In other words, our theoretical framework predicts a positive correlation between post-filler guilt and post-action guilt among separated participants. However, among unseparated participants, the post-filler guilt measure does not represent the residual level of decision guilt experienced (as these participants have not yet experienced the realization), and thus it should not correlate with post-action guilt.

To test this prediction, we regressed post-action guilt on post-filler guilt (mean-centered), condition (separated or unseparated, effect-coded), and their interaction. In line with our predictions, this analysis reveals the previously uncovered main effect of condition ($\beta = 1.93$, $t(146) = 2.82$, $p = .006$), whereby participants report higher post-action guilt in the unseparated condition than in the separated condition. There was also a main effect of post-filler guilt ($\beta = .86$, $t(146) = 5.50$, $p < .001$), whereby the higher a participant's level of post-filler guilt, the higher her reported post-action guilt (or said otherwise, the lower the level of post-filler guilt, the lower the level of post-action guilt). Further, these effects were qualified by a significant interaction ($\beta = -.67$, $t(146) = -3.03$, $p = .003$). Simple-effects analyses reveal that within the separated condition, there was the predicted positive main effect of post-filler guilt on post-action guilt ($\beta = .86$, $t(73) = 6.32$, $p < .001$). However, as expected, there was no such effect within the unseparated condition ($\beta = .19$, $t(73) = 1.08$, $p = .283$).

Discussion

Experiment 2 converges with Pilot Study 2 and Experiment 1 to demonstrate that consumers experience decision guilt that decays during a decision-enactment gap. Further, it identifies a link between this decay and the ability of decision-enactment gaps to lower the guilt consumers feel upon acting. Together, these studies establish the guilt-reducing power of decision-enactment gaps as measured via self-report. Nevertheless, there are pitfalls of emotional self-reports (e.g., Robinson and Clore 2002; Pham 2004; Winkielman

and Berridge 2004). Thus, the next experiment tests the behavioral consequences of decision-enactment gaps.

EXPERIMENT 3: DECISION-ENACTMENT GAPS INCREASE INDULGENT CONSUMPTION

Experiment 3 shifts from examining interpersonal harm to exploring self-control and consumption. Participants choose between a healthy and an unhealthy food item (Shiv and Fedorikhin 1999), and then consume their chosen item. Drawing on prior research establishing that guilt compels consumers to stop their guilt-inducing behavior (Amodio et al. 2007; Giner-Sorolla 2001; Monteith et al. 2002), we predict that decision-enactment gaps will disrupt this "stopping mechanism," enabling individuals to carry out their guilt-inducing behavior more freely. Accordingly, decision-enactment gaps should increase consumption of the guilt-inducing (unhealthy) item, but not impact the healthy item that should not elicit guilt.

Method

Overview. Two hundred fifty-one undergraduates (59% female; M_{age} = 21.0 years, SD_{age} = 2.9 years) participated for partial course credit. Participants were randomly assigned to one of two conditions (separated or unseparated), between-subjects. All participants completed five activities in this experiment. Separated participants chose a food option, completed filler tasks, reported their anticipation, consumed their chosen

food, and finally had their food weighed. Thus, there was a decision-enactment gap between their choice and consumption. Unseparated participants completed the same activities, but instead completed the filler tasks before making the food choice, eliminating the temporal gap between choice and consumption. Participants in both conditions completed the same set of tasks, and ended up consuming their chosen items at approximately the same time. In this way, we aimed to prevent intertemporal considerations from muddying the choice or consumption (e.g., Read and van Leeuwen 1998).

Food Choice and Filler Tasks. Upon entering the lab, all participants indicated how hungry they were (1: Not at all hungry to 7: Very hungry) and the time since they last ate. Participants in the separated condition then received a packet informing them that at some point in the session, they would watch and evaluate a movie. They read that we were studying whether eating during a movie affects perceptions of the movie, and so they could choose one of two food options to consume during the movie. The food options were caramel-covered popcorn (an unhealthy option) or shelled edamame (a healthy, control food). The ingredients and nutritional facts for both options were printed in the packet. A pretest described in the Appendix confirms the intuition that choosing caramel-covered popcorn over edamame induces guilt.

Participants circled an option to indicate their choice, and then proceeded to complete unrelated lab filler tasks (approximately 15 minutes long; details in the Appendix). Unseparated participants completed the same two tasks, but in the reverse order (first completing filler tasks and then choosing a food item).

Reporting Anticipation. After completing these two tasks, participants received a sheet and responded to two questions: “How excited are you to watch a movie?” (movie anticipation) and “How excited are you about eating the snack?” (eating anticipation) (1: Not at all excited; 5: neutral; 9: Extremely excited). Prior research has shown that having a longer wait time before consumption can increase anticipation and consumption enjoyment (Nowlis, Mandel, and McCabe 2004), so the latter measure was designed to account for the potential influence of this non-guilt factor on amount consumed. The movie excitement measure was designed as a distractor item.

Food Consumption. Next, research assistants handed participants their chosen items in pre-weighed paper bags (180 grams of food; details about this quantity are in the Appendix), informed participants they could eat the snack during the movie, and instructed them to put on their headphones. Participants then freely consumed their food while watching a movie clip (the first 18 minutes and 20 seconds of the documentary “Free to Play”).

Food Weighing and Post-Consumption Measures. After each participant’s movie clip ended, a research assistant collected his/her snack bag and brought it to a separate room to weigh it. Participants were not made aware that their food would be, nor that it was, weighed; they were merely told the food was removed so that they could complete the remainder of the experiment without distraction. Amount consumed was converted from weight (in grams) to volume (in milliliters) to more accurately represent how much food was consumed, but this conversion does not affect the pattern of our results (e.g., the same pattern of results arise when examining calories consumed).

Next, each participant completed a follow-up survey to maintain the cover story that we were interested in whether eating influenced their perceptions of the movie (details in the Appendix). They then reported which food they chose and why they chose it, as well as how the food tasted (1: Extremely bad to 7: Extremely good or N/A: didn't eat any). Finally, they reported demographic and individual difference information (details in the Appendix).

This experiment did not directly measure participants' self-reported guilt, as its primary focus was to capture real behavior. The process of consciously reporting one's emotional state can significantly alter emotional experiences themselves (Kassam and Mendes 2013), so measuring self-reported guilt before measuring its behavioral consequences may have distorted participants' behavior (Scheier and Carver 1977; Lieberman et al. 2007). On the other hand, measuring reported guilt after participants had acted would also be problematic, as this measurement could be influenced by the very behavior under investigation: the amount consumed. (If decision-enactment gaps increase the amount consumed, this increased consumption could accordingly increase action guilt, canceling out the decay of decision guilt.)

Results

Choice. To test for possible selection effects in the form of different choices made in the two conditions, we examined the percentage of participants in each condition who chose caramel-covered popcorn. As desired, there was no significant difference between

the proportion of separated (71.32%) and unseparated (64.75%) participants choosing popcorn ($\chi^2(1, N = 251) = .96$, $p = .327$, $\varphi = .06$), suggesting that the temporal separation manipulation did not result in two different distributions of participants' choices.

Anticipation. To ensure that separated participants were not more eager to consume than were unseparated participants, as would be predicted by a savoring or anticipation account (because more time had elapsed after their choice), we analyzed participants' self-reported excitement to eat. For this and all remaining analyses, we effect-coded (1 vs. -1) each of our binary condition variables (here, condition and food) to allow for interpretation of our main effects. A linear regression with condition, food, and their interaction as predictors revealed no main effect of condition ($\beta = -.09$, $t(246) = -.65$, $p = .515$) no main effect of food ($\beta = .15$, $t(246) = .1.15$, $p = .252$) and no interaction between food and condition ($\beta = -.05$, $t(246) = -.37$, $p = .711$). These results suggest that anticipation did not drive any differences by condition on amount consumed, but to further contend with this possibility, we also directly control for eating anticipation when examining amount consumed.

Amount Consumed. Amount consumed (in milliliters) was right-skewed (skewness = 1.47), and a Shapiro-Wilk test (Shapiro and Wilk 1965) revealed a violation of normality ($W = .86$, $p < .001$). To account for this violation, we analyzed the median amount consumed using a quantile regression; the Appendix includes alternative analyses and additional details. Condition (separated or unseparated), chosen food (popcorn or edamame), and their interaction served as predictors, and eating anticipation (mentioned above; $\beta = 15.87$, $t(235) = 4.31$, $p < .001$), time of day ($\beta = 22.56$, $t(235) = 1.37$, $p = .173$),

and perceived tastiness of the food ($\beta = 20.60$, $t(235) = 3.12$, $p = .002$) served as control variables. This analysis revealed a main effect of food ($\beta = 84.10$, $t(235) = 8.90$, $p < .001$) such that participants consumed a higher volume of popcorn than edamame, and a main effect of condition ($\beta = 19.37$, $t(235) = 2.13$, $p = .034$), whereby participants consumed more food in the separated condition than in the unseparated condition. Further, these effects were qualified by a significant interaction between food and condition ($\beta = 23.96$, $t(235) = 2.59$, $p = .010$), suggesting that condition differently affected consumption of the two foods.

An analysis of the simple effects using median quantile regressions (with the same control variables as in the prior analysis) revealed the predicted effect of condition on popcorn consumed, ($\beta = 73.93$, $t(161) = 1.97$, $p = .051$). That is, separated participants consumed more popcorn (median = 273.56 ml, 95% CI (bootstrapped) [230.70, 307.40]) than unseparated participants did (median = 221.80 ml, 95% CI (bootstrapped) [185.70, 262.10]). In contrast, there was no effect of condition on the amount of edamame consumed ($\beta = -3.25$, $t(71) = -.26$, $p = .797$), as separated participants (median = 65.41 ml, 95% CI [44.56, 83.59]) and unseparated participants (median = 62.63 ml, 95% CI (bootstrapped) [46.25, 77.90]) consumed similar amounts. Figure 1.7 presents covariate-adjusted means to complement the medians presented in the current analyses (a graphical presentation of the medians, revealing the same pattern, is presented in the Appendix).

Discussion

Experiment 3 reveals that decision-enactment gaps allow individuals to carry out their otherwise guilt-inducing behavior more freely. Specifically, it finds that temporally separating choice and consumption increased the amount of indulgent food consumed, but did not affect the consumption of healthy food. Thus, in support of our theoretical account, decision-enactment gaps can mitigate the initial function of guilt: behavior interruption (Amodio et al. 2007). Experiment 4 investigates whether decision-enactment gaps can also mitigate the subsequent function of guilt: motivating reparation (e.g., Tangney et al. 1996).

EXPERIMENT 4: DECISION-ENACTMENT GAPS REDUCE ATONEMENT

Guilt, unlike other emotions, triggers reparative actions to atone for the guilt-inducing behavior (e.g., Tangney et al., 1996). Thus, one may infer how guilty someone feels by observing the extent to which she makes amends. To this end, Experiment 4 tests whether decision-enactment gaps reduce post-behavior atonement. Participants choose between completing (a) an easy task and (b) a difficult task that will earn a charitable donation. We predict that participants who choose the easy option will feel guilty about the foregone opportunity to aid charity, prompting them to subsequently atone by donating a portion of their earnings. Indeed, prior research shows that consumers who feel guilty often alleviate this guilt by donating to charity (e.g., Zemack-Rugar et al. 2015). However, this atonement should be tempered when there is a decision-enactment gap between choice and task performance.

Method

Overview. Four hundred fifty-five MTurk workers (47% female, M_{age} = 36.5 years, SD_{age} = 11.7 years) participated for a small monetary payment. Participants were randomly assigned to one of three conditions (unseparated, separated, or decision-only), between-subjects. The third condition (decision-only) was included in this experiment to assess the consequences of decision guilt alone (more details are provided next). All participants completed four activities in this experiment. Participants learned they would complete several tasks and could choose which of two to complete at some point in the experiment. After choosing a task, separated participants completed filler tasks (approximately 3-6 minutes long; details in the Appendix), then completed the task they had chosen, and finally saw an opportunity to donate potential raffle winnings to a charity. In contrast, unseparated participants completed their chosen task immediately after choosing it. Thus, these two conditions differed only in that separated participants had a temporal gap between their task choice and the completion of that task. Decision-only participants chose a task, then saw the donation opportunity before completing their chosen task and then the filler task. Thus, they had made a decision but not performed the action before having the chance to donate. This third condition was designed to further corroborate the notion that mere decisions can rouse guilt, even before the actions associated with those decisions.

Task Choice. Participants chose between two tasks: one option was to rate the quality of three one-line jokes (e.g., "Q: What did the apple say to the worm? A: You're boring me") on a 1-10 star scale, while the other involved transcribing a paragraph of text

that was scanned in at a 90-degree rotation. Accompanying the latter task description was, "If you choose this boring task, we will also donate 25 cents to the Save the Children charity on your behalf." We expected that the majority of individuals would chose the easier joke-rating task, but that they would feel a bit guilty about choosing not to aid the Save the Children charity. A pretest described in the Appendix confirms that choosing an easy task instead of a difficult, donation-generating task would be guilt-inducing, in line with our intuitions.

Donation Opportunity. Participants learned they would be entered in a raffle to win a $25 bonus and could donate any amount of it to Save the Children if they won. In this way, participants could atone for their transgression of not aiding the charity. We designed this atonement opportunity to aid the specific charity that had been affected by participants' initial decision, based on prior findings that although guilt often promotes a general tendency to be prosocial (e.g., de Hooge, Zeelenberg, and Breugelmans 2007), it does so most strongly when the beneficiary of that prosociality is the person(s) whom the individual had wronged (Cryder, Springer, and Morewedge 2012).

Additional Measures. Finally, all participants indicated whether they had heard of Save the Children before and reported how familiar they were with the charity (1: not at all to 4: a great deal). They also indicated how often they give to charity in general (1: never to 6: six or more times a year) and provided demographic information.

Results

Choice. Given that participants in all conditions had identical experimental experiences prior to making their choice, we expected no differences in the rate of choosing the easy, guilt-inducing task (rating jokes) over the donation-generating task (transcription) in each condition. Indeed, there were no differences among the proportion of separated (79.74%), unseparated (72.67%), and decision-only (75.66%) participants who chose the jokes task ($\chi^2(2, N = 455) = 2.10$, $p = .350$, $\varphi = .07$).

Participants Choosing the Hard, Donation-Generating Task. The small proportion of participants who chose the difficult task (23.95%) selected themselves into a group of individuals who should not feel guilt, and who likely have a taste for helping charities (as they were willing to endure an arduous task to earn a 25-cent donation). As such, we would expect them to also donate a larger share of their potential raffle earnings (vs. participants who chose the easy task). However, this choice should reflect their orientation toward helping charity, with no relation to guilt; as such, these individuals' actions are outside the realm of our predictions. Our data support these notions: these individuals donated substantially more of their potential raffle earnings than did the easy-task-choosers ($M = \$10.30$, $SD = 7.85$ vs. $M = \$5.11$, $SD = 5.90$; $t(453) = 7.36$, $p < .001$), and also reported donating significantly more to charity every year ($M = 3.76$, $SD = 1.56$ vs. $M = 2.98$, $SD = 1.59$; $t(453) = 4.52$, $p < .001$). We therefore focus our analyses on those participants who chose the guilt-inducing task.

Amount Donated as a Function of Separation. Figure 1.8 displays average donation amounts by condition and by task choice. For participants choosing the guilt-inducing task, we predicted that temporal separation would reduce the amount donated. To test this, we

conducted a linear regression including these participants, with two dummy variables representing condition (one for the unseparated condition, and one for the decision-only condition) as predictors, while controlling for participants' average charitable contributions ($\beta = 1.08$, $t(341) = 5.55$, $p < .001$). (Whether they had heard of the Save the Children charity was not a strong predictor, $\beta = -.06$, $t(341) = -.18$, $p = .855$, and including/excluding this covariate does not influence the main results.) This analysis revealed the predicted effect of the unseparated condition ($\beta = 1.83$, $t(342) = 2.46$, $p = .015$), whereby unseparated participants donated significantly more (least-squares M = \$6.01, SE = .54) than did separated participants (least-squares M = \$4.18, SE = .51). There was no significant effect of the decision-only condition (least-squares M = \$5.25, SE = .53; $\beta = 1.07$, $t(342) = 1.46$, $p = .145$), illustrating that these participants donated directionally but not significantly more than did separated participants.

We next conducted a similar regression in which the dummy variables instead represented the decision-only and the separated conditions, so that the unseparated condition became the reference level. This analysis revealed a non-significant effect of the decision-only condition ($\beta = -.76$, $t(342) = -1.01$, $p = .314$). Thus, participants who had decided which task to complete—but not yet completed it—fell in between the separated and unseparated participants. Although they did not significantly differ, the relative levels fell in line with expectations, whereby unseparated participants donated directionally more than did decision-only participants, who themselves donated directionally more than did separated participants. More importantly, participants for whom there was a decision-

enactment gap between task choice and task completion indeed atoned less (i.e., donated less) than did those with no temporal separation.

Discussion

Supporting the second behavioral consequence of guilt, Experiment 4 finds that decision-enactment gaps decrease post-behavior atonement in the form of charitable donations. These findings corroborate the role of guilt in our theoretical framework and identify an important consequence of decision-enactment gaps.

GENERAL DISCUSSION

Decades of research have explored how consumers choose between "oughts" and "wants," between prosocial and selfish acts, between indulgence and restraint. In the face of such conflicts, and against our better nature, we often neglect our workout routines, turn down charity appeals, break our diets, and spend more than we should. These actions, although often inherently pleasurable, can also elicit the unpleasant feeling of guilt. The current research demonstrates that merely separating in time an initial decision to perform such behaviors from their actual performance can attenuate this emotion, reducing how guilty individuals feel while acting. These "decision-enactment gaps" can thus lead us to indulge with less guilt (Experiments 1 and 2), consume more than we otherwise would (Experiment 3), and atone less for our behavior afterward (Experiment 4).

This research also introduces the notion that mere realizations that one will act can elicit guilt. This "decision guilt" is distinct from the guilt consumers anticipate experiencing upon acting (Pilot Study 1), and importantly, decays over time (Pilot Study 2), in the period between decisions and actions. This decay contributes to the reduction of overall guilt felt while acting (Experiment 2). Furthermore, the reduction in guilt resulting from decision-enactment gaps arises both when individuals freely make a choice to perform a guilt-inducing act (Experiments 3 and 4) and when they do not freely decide, but merely perceive a sense of personal responsibility in determining that they will perform the action (Pilot Studies 1 and 2, Experiments 1 and 2).

In addition to offering evidence for our predicted effects, the current data also address certain alternative explanations. First is the possibility that other emotions, such as regret or shame, might be responsible for the current results, rather than guilt. Although many emotions decay over time (Ekman 1984; Gilovich et al. 1998; Jager et al. 2014; Kahneman 1995), Experiment 1 provides evidence that only guilt (and not shame, regret, sadness, (un)happiness, or fear) decayed in the period between an initial realization that one would act and its actual enactment. Further, the behavioral responses to decision-enactment gaps identified in Eperiments 3 and 4—increased indulgence and reduced atonement—align particularly with the dual functions of guilt: interruption and reparation (Amodio et al. 2007). Together, these results underscore the notion that guilt is acutely affected by decision-enactment gaps.

An additional possible concern is the potential for selection effects to influence our findings. For example, participants in Experiment 3 chose between eating caramel-covered

popcorn (an unhealthy, though admittedly delicious, indulgence) and edamame (a healthy food) either toward the beginning of the experimental session (i.e., the "separated" condition) or after completing a few minutes of filler tasks (i.e., the "unseparated" condition). One might speculate that intertemporal considerations led individuals to select their food options differently in the two conditions. However, participants were not aware of when they would be consuming the food items, and indeed rates of choice were equivalent in both conditions. This pattern repeated in Experiment 4, as well. Furthermore, we find the predicted effects of decision-enactment gaps even when participants did not choose their actions at all (Pilot Studies 1 and 2, Experiments 1 and 2). Thus, selection effects cannot explain these results. Instead, these findings support the notion of guilt decaying during decision-enactment gaps.

Theoretical Contributions

The present work makes notable contributions to the consumer behavior literature. First, it provides a deeper understanding of the consequences of what we term decision-enactment gaps: the time that passes between decisions and actions. Past research has investigated how people choose different actions for the present versus the future; in other words, this tests how knowing that a decision-enactment gap will arise influences choice. For example, individuals tend to choose healthier items to consume in the future than in the present (Read and van Leeuwen 1998; VanEpps, Downs, and Loewenstein 2016). Accordingly, such work often touts the benefits of "precommitting" by reaching decisions

earlier (but see Kivetz and Simonson 2002), such as by ordering meals in advance. Our research complements these findings in two ways. First, it explores the effects of decision-enactment gaps not on what is chosen, but on the actual experience of the chosen behavior. Second, it notes an important caveat to the notion that consumers behave more virtuously when choosing for the future. Although individuals may order fewer unhealthy treats in advance (vs. shortly before consumption), the items they do pre-order are likely to elicit less guilt while eating. Paradoxically, then, consumers might restrict their in-the-moment consumption of these unhealthy items more when their order immediately precedes consumption (as in Experiment 3), undermining some benefits of pre-ordering. Future research may seek to identify the conditions under which delays are (vs. are not) beneficial for consumers' overall well-being.

This work also builds on prior research examining non-guilt-related effects of decision-enactment gaps on the experience of behaving, an understudied stage in the consumption process (Pham 2013). One effect of such gaps is that they may allow for savoring, which is the current pleasure derived from anticipating a future consumption experience (Amir and Ariely 2007; Chun, Diehl, and MacInnis 2017; Nowlis, Mandel, and McCabe 2004). Savoring shares certain conceptual features with decision guilt. However, savoring and anticipation tend to build up over time (Amir and Ariely 2007), unlike decision guilt, which decays. Moreover, savoring in a period prior to behavior tends to increase ongoing enjoyment during the experience (Chun, Diehl, and MacInnis 2017). In contrast, we find that the decay of decision guilt results in lower guilt during an experience. Guilt is also different in this respect from anxiety or dread, which may

intensify over time, and which may also amplify negative emotional experiences (consider the familiar mental anguish of dreading going to the dentist, and how that buildup makes you feel once you get to the office). This suggests a sharp contrast between emotions that increase (vs. decrease) over time and amplify (vs. attenuate) the complementary emotional experiences during an action—a ripe opportunity for future research.

The current research also draws attention to the dynamic nature of emotions in consumer behavior. Whereas a rich body of literature has examined the antecedents and consequences of guilt and other emotions in consumer contexts (e.g., Allard and White 2015; Cavanaugh, Bettman, and Luce 2015; Coleman et al. 2017; Morales, Wu, and Fitzsimons 2012; Zemack-Rugar et al. 2007), much of this work takes a "snapshot" view of emotions, studying them at one point in time. By positing a dynamic model of guilt and tracing its implications not only over time, but also through actual consumption, this research improves our understanding of the roles of specific, functional emotions while highlighting the benefits of developing a richer temporal framework for them.

Limitations and Directions for Future Research

Our theoretical framework, and the limitations of the current work, provide opportunities for future research. When consumers reach decisions, our results suggest that their corresponding guilt can decay prior to acting—but what is responsible for this decay? The introduction presents a set of candidate explanations; however, which account can best explain this decay remains an open question for future research.

Regardless of the cause of this decay, we find that decision-enactment gaps have a measurable impact on how guilty consumers feel when acting. Yet, another key question arises: what do decision-enactment gaps do to the total amount of guilt experienced across an entire decision-enactment episode? Decision-enactment gaps may simply shift the experience of decision guilt to an earlier period, retaining the same total amount of guilt experienced. An alternate possibility is that decision guilt and action guilt interact when they are experienced in close temporal proximity (i.e., when there is no decision-enactment gap): perhaps there is a diminishing marginal impact of guilt, so that the total amount of guilt experienced concurrently is lower than when these two sources are experienced separately (sub-additivity). A similar attenuation effect could arise if guilt interacts with other emotions in the course of behaving (e.g., the joy of indulging) in such a way that these emotions reduce each other's impact, or even amplify one another (Goldsmith, Cho, and Dhar 2012). Determining the total guilt consumers experience is an important future question from a consumer welfare perspective.

Furthermore, future research might investigate whether decision-enactment gaps influence remembered guilt a period of time after the behavior. Might the reduction in guilt experienced during the behavior influence the amount of guilt consumers remember having experienced? Prior research gives reason to believe that changing consumers' guilt during experiences may impact the guilt they recall even weeks afterward (Goldsmith et al. 2012). If decision-enactment gaps reduce not only experienced guilt but also remembered guilt, this effect may have additional significant consequences, as individuals often rely on their memory of past emotional experiences to guide their decisions (Baumeister et al. 2007).

There exist additional interesting directions for future research in exploring the behavioral consequences of decision-enactment gaps. For example, through reducing consumers' guilt, decision-enactment gaps may decrease the extent to which individuals opt for utilitarian over hedonic items post-consumption (Zemack-Rugar et al. 2007; Ramanathan and Williams 2007; Levav and McGraw 2009). In this way, decision-enactment gaps may encourage individuals to continue indulging after their initial indulgences, producing a consumption cycle.

Our experiments all involved situations in which participants made a choice (or, were partly responsible for the decision) to engage in a particular action and then carried out that action; there was no opportunity for them to change their minds prior to the behavior. This is true of many situations in consumers' lives. However, in some cases, consumers reach decisions to engage in an action, but then may re-think their choice. Experiencing decision guilt in advance of an action may, in certain circumstances, motivate consumers not to perform the action after all. Further research is needed to identify when this pattern could arise.

Finally, the current research focused specifically on guilt because of the significant consequences it has for consumer behavior. However, one might question whether a similar dynamic exists for other emotions. Our framework takes into account two key features of guilt: (a) its associated sharpened focus on a specific behavior, rather than a broader evaluation of the self (Lewis 1971; Tracy and Robbins 2004), and (b) its attribution of personal responsibility and self-blame for the behavior (Smith and Ellsworth 1985; Tangney et al. 1996). That is, guilt follows from a cognitive realization that one will

cause harm that is one's own responsibility—an awareness that can plausibly develop after merely reaching a decision or coming to a realization, thus laying the foundation for decision guilt. The prediction that decision-enactment gaps will lower guilt during an action follows from both the existence of this decision guilt and its decay over time. Although many other emotions also decay over time, it is less clear whether such emotions would have a similar decision component. Consider the emotion of shame, which involves an attribution of personal responsibility, like guilt: although shame may decay over time (as in Pilot Study 2), it tends not to be associated with a focal choice or realization, but rather a global evaluation of oneself/one's identity (Tangney et al. 1996). For this reason, it is possible but less likely for decision-enactment gaps to reduce the experience of shame (indeed, we found no such effect in Experiment 1), although further research is needed to investigate whether and when decision-enactment gaps might impact shame and other emotions. Note that our measures for shame only included self-reports; to properly investigate the dynamic effects of shame, one ought to design behavioral experiments that focus on this emotion.

Practical Implications

In the marketplace, consumers engage in numerous actions with the potential to induce guilt. These findings have broad application, given that decisions and their enactment are naturally separated in many situations in individuals' lives. Students may purchase vending machine snacks before class to consume afterward, and may decide

either in advance or mid-viewership how many episodes of a TV show to watch on a school night. Customers can use mobile apps to place advance orders for restaurant meals (e.g., Zest) and cappuccinos (e.g., Starbucks), and can pre-order video games (e.g., at Best Buy), movies, and more. Our findings suggest that these pre-ordering decisions can influence the emotional costs consumers bear during these experiences.

Based on this notion, marketers of products and services that may elicit guilt, such as calorie-dense desserts, extravagant vacations, or luxurious massages, may wish to encourage consumers to reach the decision to indulge further in advance. Sending advance mailings to encourage consumers to "precommit" to indulging may not only increase their likelihood of actually doing so (e.g., Gollwitzer 1999), but may also improve their experiences by mitigating the unpleasant experience of guilt. From a hedonic perspective, consumers may be better off in terms of their enjoyment when including decision-enactment gaps.

On the other hand, the reduction of guilt associated with decision-enactment gaps may have adverse consequences that marketers and policymakers wish to prevent. Guilt performs the beneficial functions of prompting behavioral interruption and atonement. Thus, individuals who indulge, perhaps by consuming an unhealthy dessert or buying an expensive vacation, may be less inclined to go to the gym or restrict future expenses to make up for these actions with (vs. without) a decision-enactment gap. Marketers of products and services that can alleviate guilt, or policymakers seeking to encourage beneficial post-behavior atonement, may wish to caution against pre-commitment to indulgence.

Conclusion

Just as financially pre-paying for experiences prevents pain-of-payment considerations from undermining enjoyment (Prelec and Loewenstein 1998), and just as individuals adapt to prior financial costs with the passage of time (Gourville and Soman 1998), this article finds that individuals can "pre-pay" their guilt, adapting to these emotional costs prior to the commencement of their chosen experiences. Thus, adding time between a decision and its enactment can protect experiences from the pain of guilt. Conversely, our findings also suggest that reducing temporal separation between individuals' choices and their enactment can aid self-control efforts and enable adaptive guilt mechanisms to better manage behavioral challenges.

ACKNOWLEDGEMENTS

We thank Dan Ariely, Lisa Cavanaugh, Linda Hagen, Jon Jachimowicz, Minah Jung, Jonathan Levav, Alicea Lieberman, Wendy Liu, Andrea Morales, Chris Oveis, Yuval Rottenstreich, Itamar Simonson, Leaf Van Boven, Kathleen Vohs, and Katherine White for helpful comments on earlier versions of this paper. We are also grateful to the editors, associate editor, and reviewers of the published paper for their invaluable input. We also thank Jyoti Jhita, Yael Horwitz, and the Atkinson Behavioral Lab for data collection assistance.

Chapter 1, in full, is a reprint of previously published material as it appears in the Journal of Consumer Research, 45 (6): 1254—1273, 2019, Duke, Kristen Elizabeth and On Amir. The dissertation author was the primary investigator and author of this paper.

FIGURES

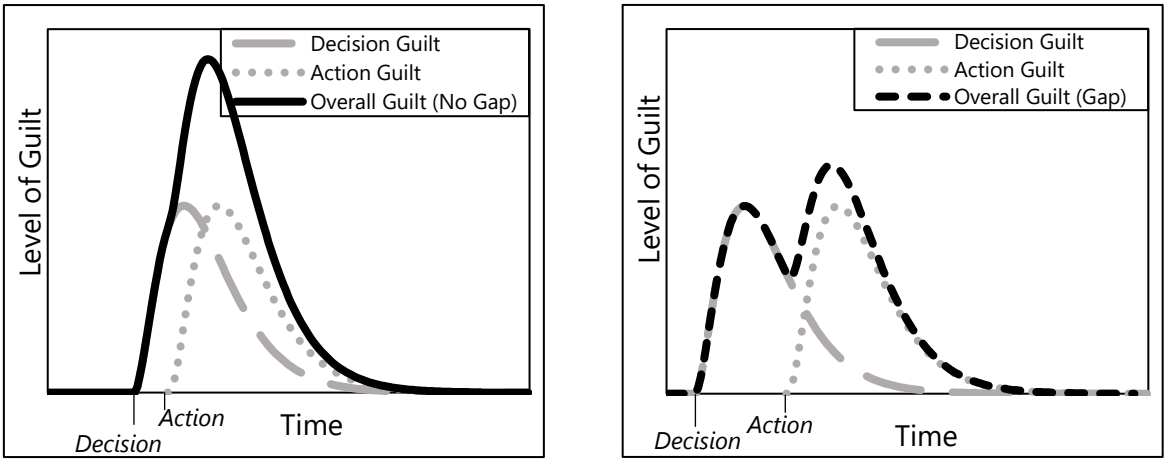


Figure 1.1. Dynamics of guilt with and without decision-enactment gaps. The left panel represents guilt without a decision-enactment gap. The right panel represents guilt with a decision-enactment gap.

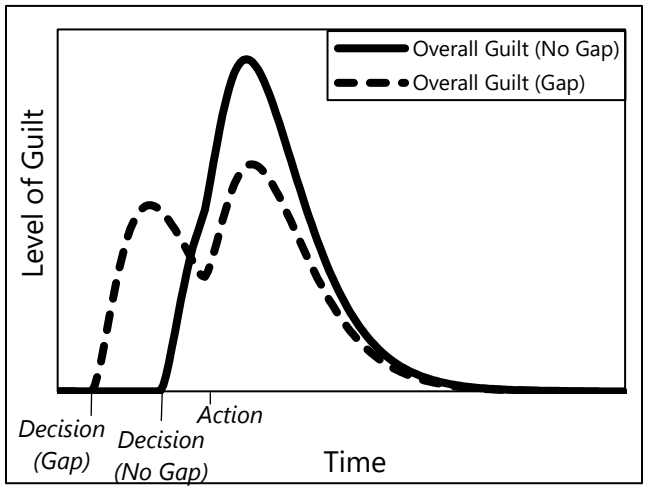


Figure 1.2. Comparing predicted overall guilt with and without a decision-enactment gap.

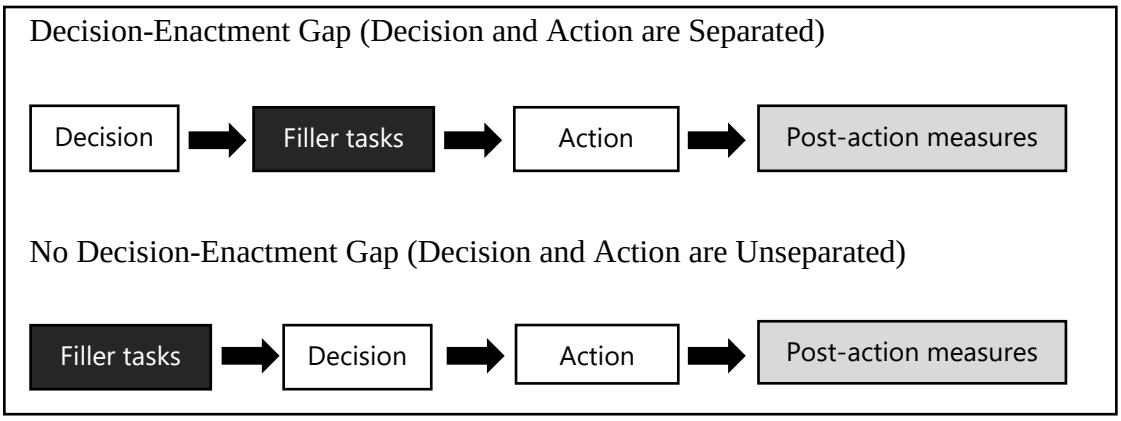


Figure 1.3. General overview of experimental timelines, Experiments 1—4.

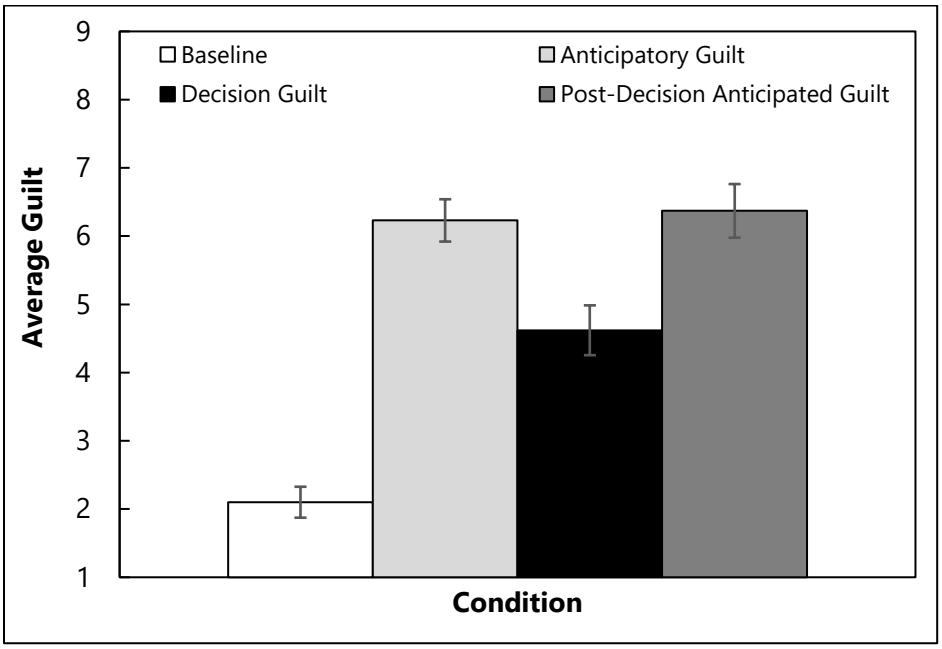


Figure 1.4. Current or anticipated guilt in each condition, Pilot Study 1. Error bars represent ±1 SE around the mean.

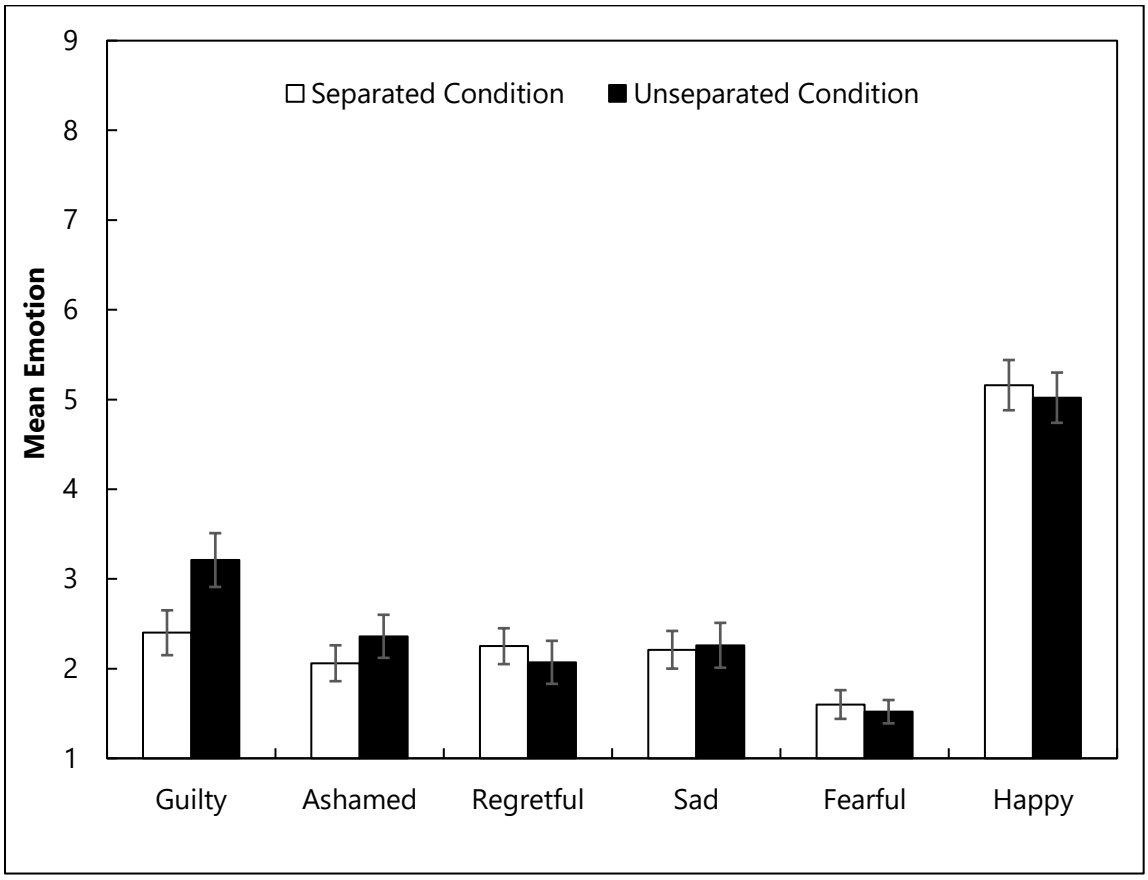


Figure 1.5. Self-reported emotions, Experiment 1. Error bars represent ±1 SE around the mean.

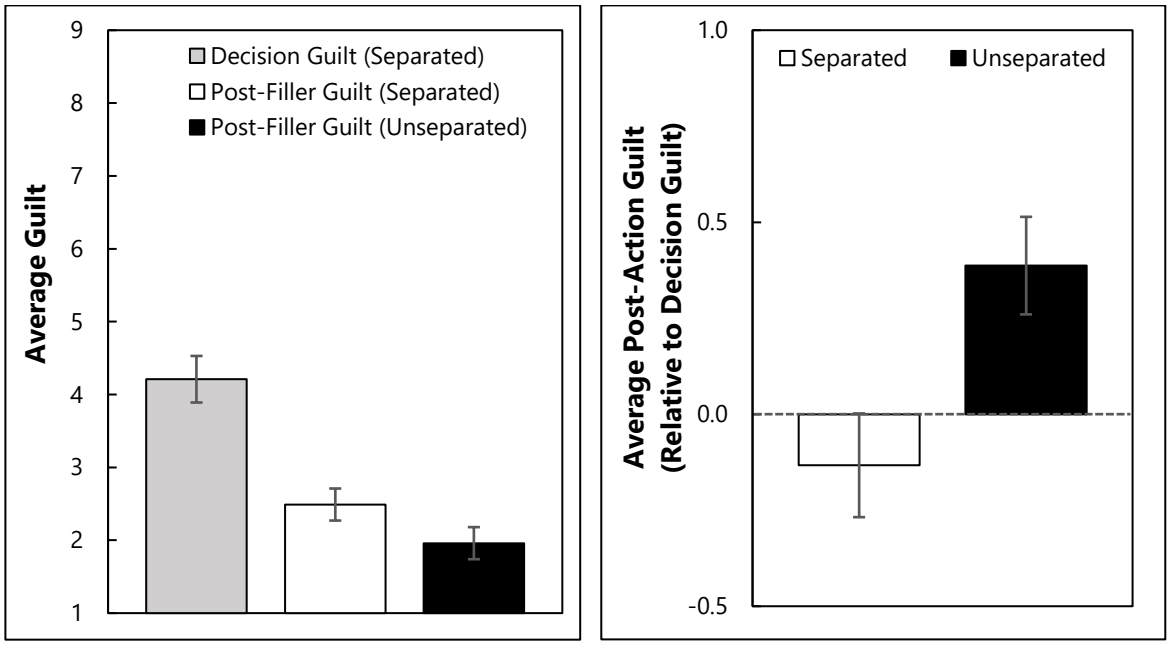


Figure 1.6. The decay of decision guilt, Experiment 2. In the left panel, the left and middle bars show the level of guilt felt at the point of the decision and the guilt felt after a delay, demonstrating the decay of decision guilt (in the separated condition, where this decay was possible). Comparing the middle and right bars reveals that the guilt felt after a delay remained higher than the guilt felt after the same filler tasks, but absent a decision (i.e., the unseparated condition). Error bars represent ±1 SE around the mean.

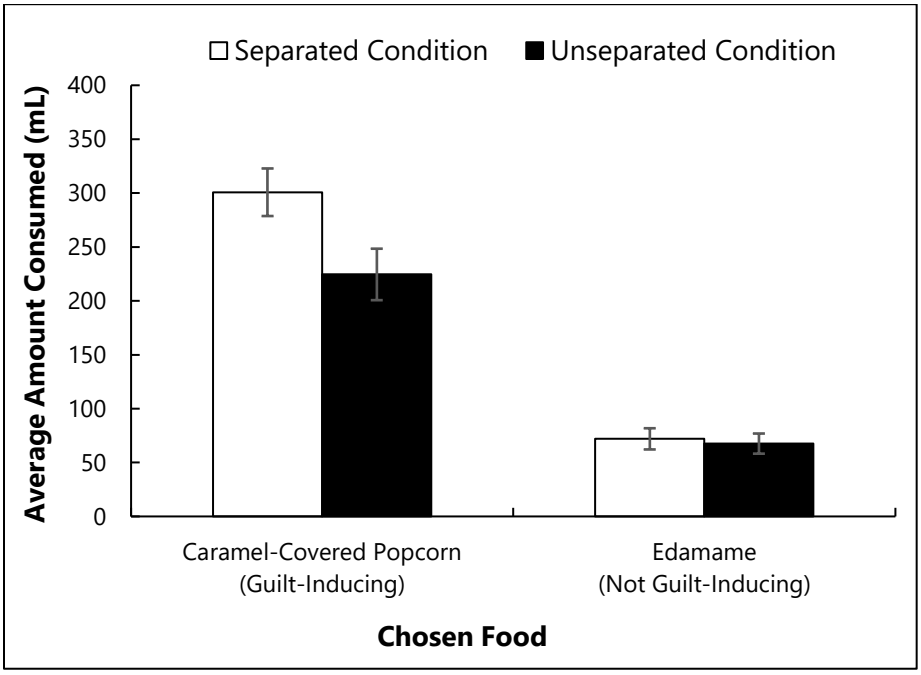


Figure 1.7. Covariate-adjusted amount of food consumed, Experiment 3. Outliers were excluded based on the Median Absolute Deviation method (see Appendix for details). The covariates were time of day, eating anticipation, and perceived tastiness of the food. Error bars represent ±1 SE around the mean.

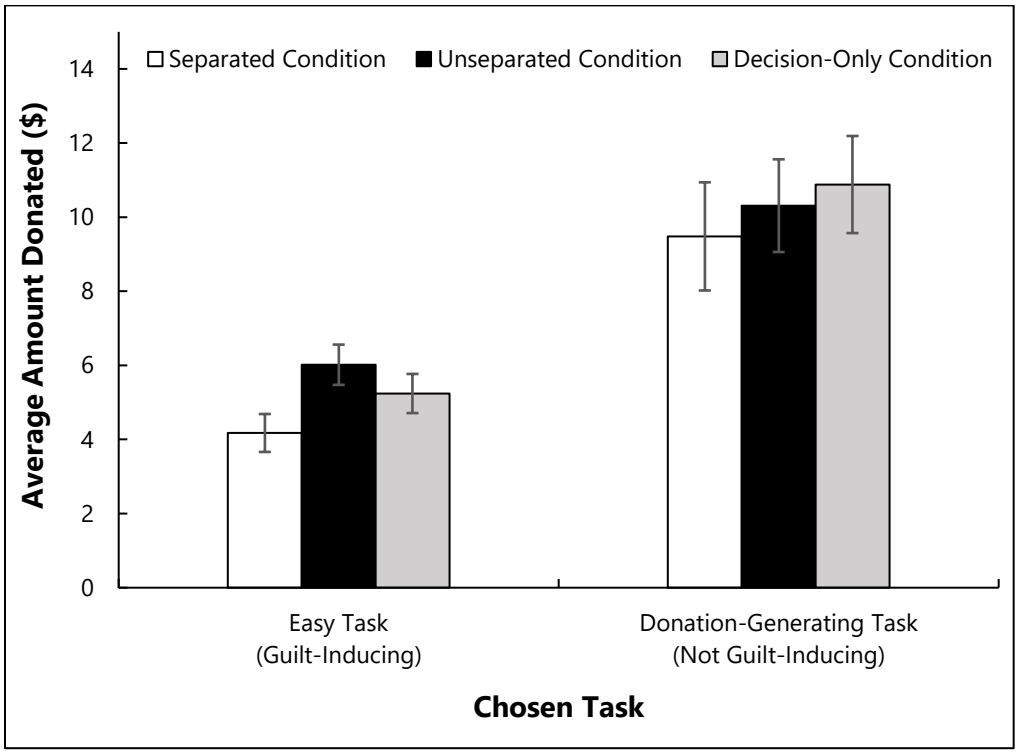


Figure 1.8. Covariate-adjusted amount donated, Experiment 4. The covariates were awareness of the charity and average annual charity contributions. Error bars represent ±1 SE around the mean.

APPENDIX

Pilot Study 1

Sample Size Determination

Two hundred forty-two undergraduates (59% female, M_{age} = 21.2 years, SD_{age} = .6 years) participated for partial course credit. We determined sample size by aiming to collect 50 participants per condition across 4 conditions. (Our lab runs studies by either the day or the week; we elected to run for three days to achieve an estimated $N \geq 200$). This experiment required no exclusions.

Attention Check

At the beginning of the experiment, participants read the following:

THE GAME OF RICH AND POOR

Today, you will be playing a game alongside your fellow lab participants from either the current session or a past session.

The decisions you make will affect **how much money you can win**, and how much money **other participants can win.**

Choose wisely!

On the next page, they responded to the question:

Quick comprehension check:

What can **your decisions** affect?

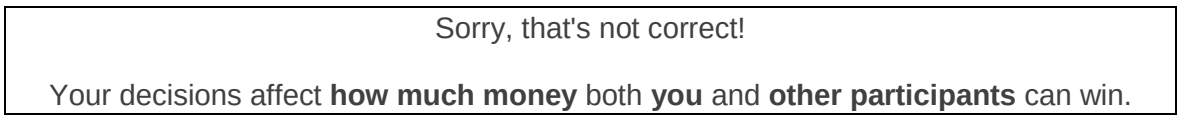
- ○ Just how much money I can win
- ○ Just how much money other participants can win
- ○ How much money both I and other participants can win

Participants who got the question correct saw the following on the next page:

That's correct!

Your decisions affect **how much money** both **you** and **other participants** can win.

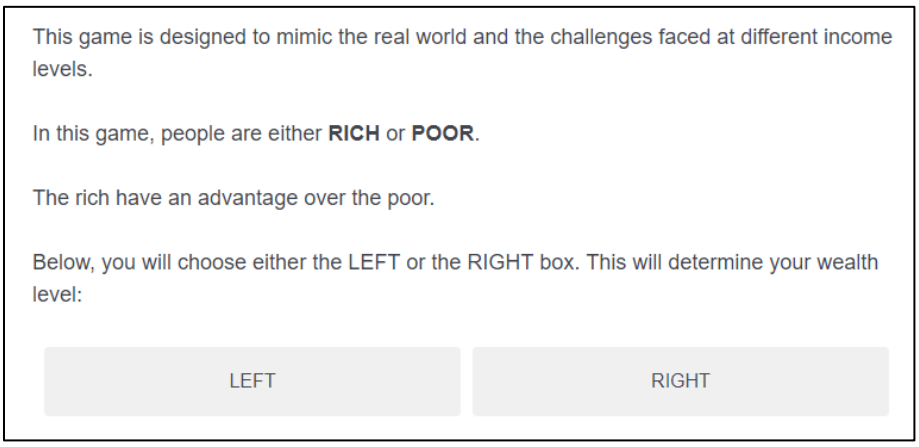
Participants who got the question wrong instead saw the following on the next page:



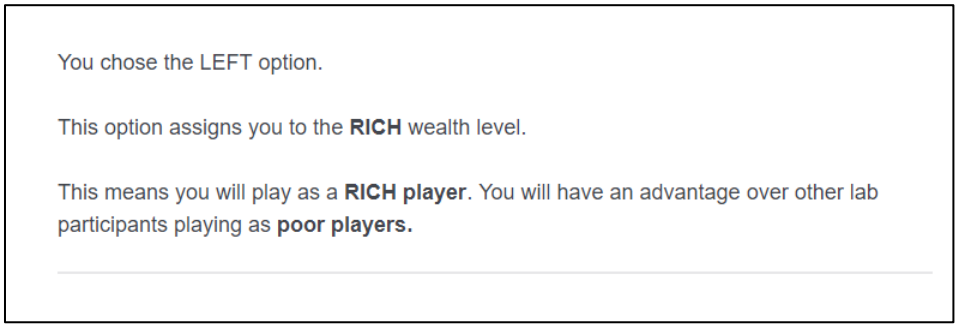
Thus, this check was designed to reinforce the importance of participants' choices and to ensure they would read task materials. We do not exclude participants based on their responses to this check in any experiment, although doing so yields the same patterns of results (and in some cases, stronger results).

Full Text of Game Situations and Choices

After the attention check, participants proceeded to the next screen and chose an option:



The following page read (if they chose the "right" option, this was reflected in the text):



The next pages described the game:

THE GAME OF RICH AND POOR

In this game, the rich people have the power to act in many ways:

- Some rich people steal from the poor to make themselves even richer.

- Other rich people instead donate their own money to help the poor.

Poor people have to work very hard. But sometimes, they get lucky and receive an **inheritance** of money.

The game will work as follows:

Lab participants playing as **poor people** will work hard to complete a challenging task.

Some of these poor people will get lucky and receive an **inheritance** ($50).

At the end of the week, one participant will be selected to receive any money won in this game.

Rich people will do one of the following:

- Rob a poor person (and steal the inheritance bonus, if that poor person had one), or
- Donate to a poor person.

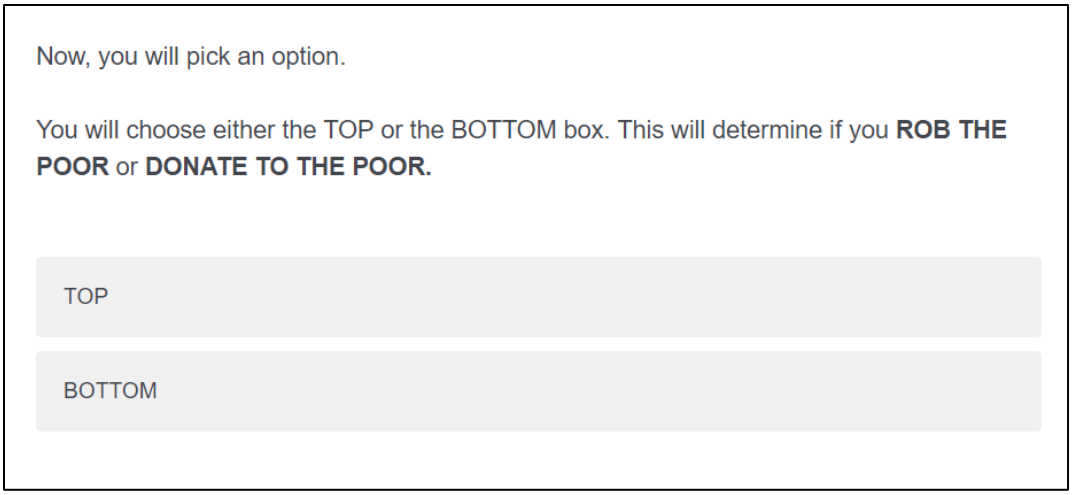
Based on your selection, you are playing as a **rich player**.

This means that you will choose between two options in the same way you chose between rich and poor.

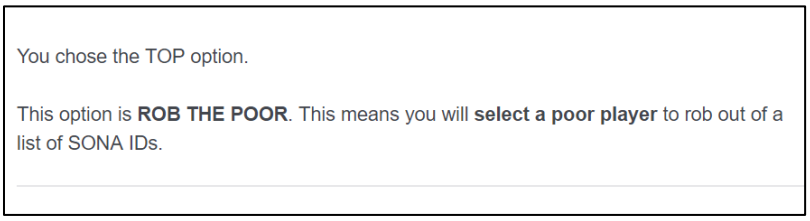
If the option you choose is **ROB THE POOR**, you will select a poor player. If that player earned an inheritance ($50), you get to keep it, and that player gets nothing.

If the option you choose is **DONATE TO THE POOR**, you will receive a $25 bonus. You will also select a poor player and give an additional $25 to that player.

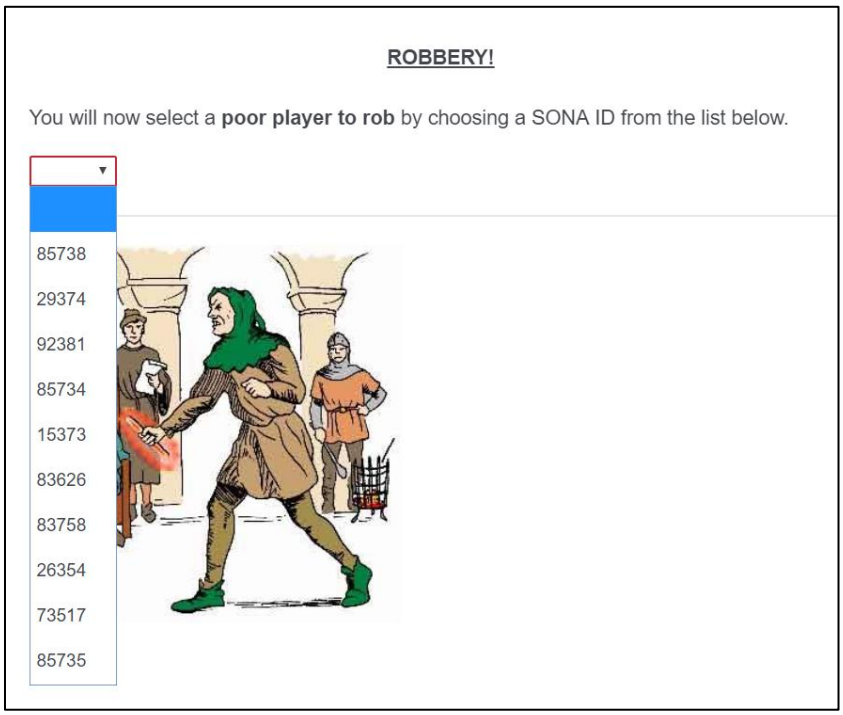
Following, participants made a choice:



The next page read (if they chose the "right" option, this was reflected in the text):



Finally, the actual theft showed the following screen. Participants needed to click the drop-down menu and select a SONA ID:



Pilot Study 2

Sample Size Determination

One hundred ninety-nine Mechanical Turk workers (51% female, M_{age} = 36.9 years, SD_{age} = 11.8 years) participated. We determined sample size by aiming to collect 100 participants per condition across 2 conditions. This experiment required no exclusions.

Filler Task Details

In the filler task, participants reviewed a choice between two possible gambles involving gains, and provided descriptions of each gamble option. They then dragged a set of eight numeric sliders (0-100) to the "67" value.

Experiment 1

Sample Size Determination and Participant Exclusions

One hundred twenty-seven undergraduates were recruited to participate for course credit. We determined sample size by a predetermined lab session count, collecting data

for one full week in our laboratory across five sessions per day (all available slots in our lab). Six participants were excluded from this study for violating the study protocol. One individual took control of the computer and clicked back to the survey instructions when she was supposed to be describing the game, one participant skipped the screen indicating what he/she should describe and the research assistant attempted to ad-lib her own instructions, two participants skipped over the first audio recording opportunity entirely, one participant requested that the research assistant read instructions out loud for him/her before recording, one participant controlled his own audio recordings (which were supposed to be done by the research assistant), and one participant did not speak at all during the first recording. Because of these deviations from the study protocol, we excluded these participants, although including them does not change the pattern of results. Thus, our final sample included one hundred twenty-one undergraduates (33% female; M_{age} = 21.5 years, SD_{age} = 1.9 years).

Filler Task Details

In the filler task, participants counted the number of occurrences of the letter "e" in two paragraphs of text, and then dragged a set of eight numeric sliders to the "67" value.

Experiment 2

Sample Size Determination

One hundred fifty undergraduates (60% female, M_{age} = 21.2 years, SD_{age} = 2.7 years) participated for partial course credit. We determined sample size by aiming to collect 50 participants per condition across 2 conditions. (Our lab runs studies by either the day or the week; we elected to run for two days to achieve an estimated $N \geq 100$). This experiment required no exclusions.

Filler Task Details

In the filler task, participants generated the first word that came to mind in response to 10 neutral words (chin, stiff, lamp, garment, avenue, spray, custom, elbow, whistle, hammer). They then solved eight simple math problems (e.g., 5*7 + 2 = ?).

Experiment 3

Pretest

We conducted a pretest to verify that individuals in our experimental population perceived the unhealthy item (caramel-covered popcorn) to be guilt-inducing, in line with our intuitions. Three hundred eighty-four undergraduates (demographics not collected) drawn from the same population as the main study saw the materials for the main experiment as part of another study session in the laboratory. They saw the same top sheet, nutritional information, and choice sheet as participants in the main study. They then answered, "Now, we would like you to imagine you were asked to choose between these food options. Which food would you like?" with choice options "Caramel-covered popcorn" and "Shelled edamame." On the following page, they answered, "You chose [selected item] instead of [unselected item]. To what extent would you feel guilty?" (1: Not at all to 9: Very much). As intended, participants who chose the caramel-covered popcorn expected to feel significantly guiltier ($M = 2.88$, $SD = 2.24$) than those who chose shelled edamame ($M = 1.60$, $SD = 1.41$; $t(382) = 6.23$, $p < .001$). Thus, participants shared our intuition that the caramel-covered popcorn could induce guilt.

Sample Size Determination and Participant Exclusions

Two hundred fifty-five undergraduates were recruited to participate for course credit. We determined sample size by a predetermined lab session count, collecting data for two full weeks in our laboratory across two sessions per day (at 2:00 PM and 3:30 PM). Four observations were excluded from analyses: three participants were accidentally given the incorrect food by research assistants, and one participant accidentally participated twice, so her second participation record was removed. Thus, the final sample included 251 undergraduates (59% female; $M_{age} = 21.0$ years, $SD_{age} = 2.9$ years).

Filler Task Details

The filler task was a set of two experiments for a separate researcher that were unrelated to the present investigation. In the first, participants completed a coordination game (e.g., Schelling 1960) for possible monetary compensation. In the second, they completed a risky gamble choice for possible monetary compensation.

Food Weight Determination

All participants in both conditions received 180 grams of their chosen snack. We set this value to be at least twice as large as the serving size for each food item, to mitigate potential ceiling effects; the serving size for the edamame was 85 grams, and the serving size for the caramel-covered popcorn was 64 grams. Note that we equated weight (rather than volume) as this was easier for research assistants to measure, although it does not represent equally matched amounts in terms of caloric content or density (one bite of edamame is approximately equivalent to eating five bites of popcorn).

Cover Story Items and Additional Measures

As mentioned in the main paper, each participant completed a follow-up survey (via Qualtrics) after watching the movie. Participants reported how good the movie was (1-10 stars), how much they enjoyed the movie-watching experience (0-100), and how engaged they were while watching the movie (1: Extremely bored to 7: Extremely engaged). None of these measures significantly differed by condition or by food chosen (p's $> .49$).

Participants also responded to additional exploratory measures: in general, how much they like edamame and caramel-covered popcorn (each on a sliding scale from 0-100), and how healthy they perceive edamame and caramel-covered popcorn to be (1: Very unhealthy to 7: Very healthy). Participants also indicated whether they were currently dieting or were on a diet at some point within the past year, and reported agreement with two statements: "Eating healthy, wholesome foods is important to me," and "I care about what I eat because it can affect my appearance" (1: Strongly disagree to 7: Strongly agree). Next, they indicated how often they tend to feel guilty about eating indulgent, unhealthy foods (1: Never to 7: Always). None of these factors were related to the amount of food participants consumed (our dependent variable of interest), and so were not analyzed further.

Consumption: Alternate Analyses

As a robustness check on the results presented in the main paper, we examined volume consumed using the Median Absolute Deviation (MAD) method for outlier detection, as is recommended practice in the presence of potential outliers (Leys et al. 2013). In this method, outliers are considered values more extreme than the median ± 2.5 times the MAD (Leys et al. 2013). In light of the difference in volume consumed between edamame and popcorn, we computed outliers within each food; this method excluded 9 popcorn participants and 4 edamame participants (note that computing outliers across foods would lead to a larger number of exclusions). On this data, we analyzed volume consumed using a linear regression with the same covariates as in the analysis in the paper (time of day, eating anticipation, and perceived tastiness of the food), and chosen food, condition, and their interaction as predictors (with the binary variables effects-coded, as in the paper). This analysis revealed a main effect of food ($\beta = 89.37$, $t(222) = 8.58$, $p < .001$), again demonstrating that participants consumed a higher volume of popcorn than edamame, and a marginal main effect of condition ($\beta = -17.32$, $t(222) = -1.72$, $p = .087$), qualified by a significant interaction between the two ($\beta = -20.07$, $t(222) = 1.99$, $p = .048$). This pattern mirrors the results presented in the main paper.

As in the paper, we unpack this interaction in two analyses. A linear regression among participants who chose popcorn revealed a significant effect of condition on volume consumed ($\beta = 76.24$, $t(152) = 2.92$, $p = .004$). Participants in the separated condition consumed significantly more popcorn (adj. M = 300.77 ml, SE = 22.07 ml) than did participants in the unseparated condition (adj. M = 224.53 ml, SE = 23.88 ml). In contrast,

temporal separation did not affect the quantity consumed of edamame, the non-guilt-inducing food: a linear regression revealed no effect of condition (β = -4.42, t(67) = -.40, p = .693; separated: adj. M = 71.98 ml, SE = 9.84 ml; unseparated: adj. M = 67.57 ml, SE = 9.34 ml). These results converge with those presented in the main paper.

Graphical Display of Median Amounts Consumed

In the main paper, we chose to present the adjusted-mean amount of food consumed in each condition (as calculated in the analyses above) to allow for a more familiar presentation of the results (i.e., including error bars). In figure A1.1, we present a graph of the median amount of food consumed in each condition for a fuller picture of the data. As mentioned in the paper, these patterns of results are quite similar and reveal the same story.

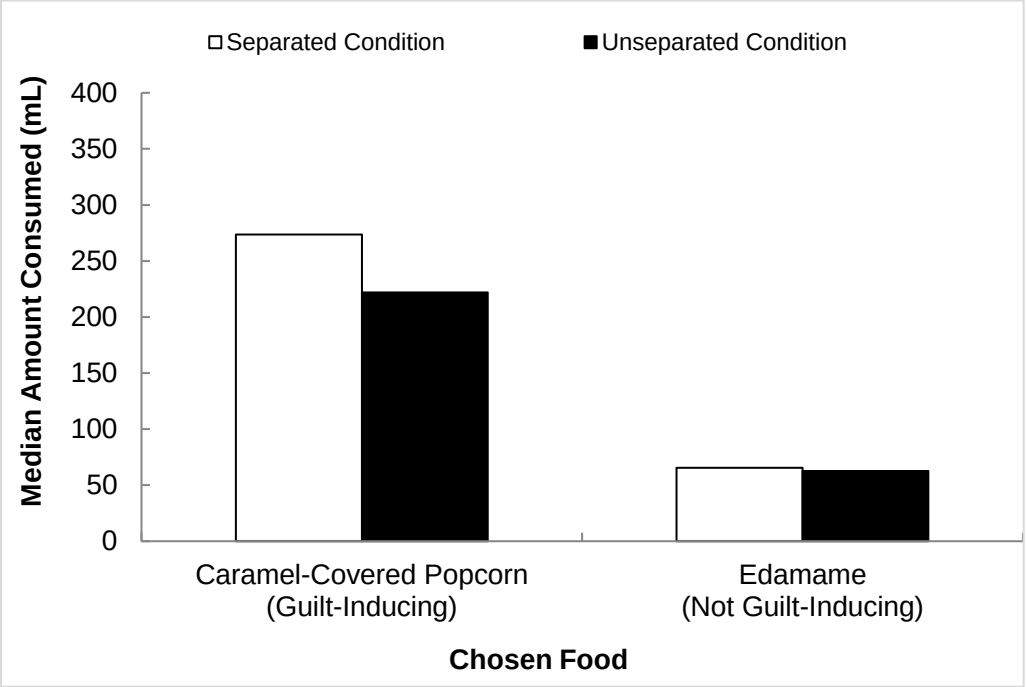


Figure A1.1. Median Amount Consumed in Each Condition, Experiment 3.

Experiment 4

Pretest

We conducted a pretest to verify that individuals in our experimental population thought that choosing an easy task instead of a difficult, donation-generating task would be guilt-inducing, in line with our intuitions. One hundred Mechanical Turk workers (41%

female, M_{age} = 33.7 years, SD_{age} = 9.9 years) imagined a participant facing a choice between an easy task and a donation-generating task (described in the same way as in the full experiment). Half of these participants imagined that the person in the scenario chose the easy task, while the other half imagined that the person chose the donation-generating task. All participants then indicated how guilty they thought the person in the scenario would feel (1: Not guilty at all to 7: Extremely guilty). As intended, participants reading about the easy-task-chooser expected this person to feel significantly guiltier (M = 2.69, SD = 1.66) than those reading about the donation-generating-task-chooser (M = 1.88, SD = 1.62; $t(98) = 2.47$, $p = .015$). Thus, participants shared our intuition that choosing the easy task could induce guilt.

Sample Size Determination

Four hundred fifty-five Mechanical Turk workers (47% female, M_{age} = 36.5 years, SD_{age} = 11.7 years) participated for a small payment. We determined sample size by intending to collect 150 participants per condition across 3 conditions (5 individuals completed the survey without submitting for payment). This experiment required no exclusions.

Filler Task Details

In the filler task, participants calculated sets of probabilities (e.g., the percentage likelihood that a six-sided die will come up “1”) and unscrambled a set of neutral sentences.

REFERENCES

Allard, Thomas and Katherine White (2015), “Cross-Domain Effects of Guilt on Desire for Self-Improvement Products,” *Journal of Consumer Research,* 42 (3), 401–19.

Amir, On and Dan Ariely (2007), “Decisions by Rules: The Case of Unwillingness to Pay for Beneficial Delays,” *Journal of Marketing Research*, 44 (1), 142—52.

Amodio, David M., Patricia G. Devine, and Eddie Harmon-Jones (2007), “A Dynamic Model of Guilt: Implications for Motivation and Self-Regulation in the Context of Prejudice,” *Psychological Science,* 18 (6), 524–30.

Baumeister, Roy F., Arlene M. Stillwell, and Todd F. Heatherton (1994), “Guilt: An Interpersonal Approach,” *Psychological Bulletin,* 115 (2), 243–67.

Baumeister, Roy F., Kathleen D. Vohs, C. Nathan DeWall, and Liqing Zhang (2007), “How Emotion Shapes Behavior: Feedback, Anticipation, and Reflection, Rather than Direct Causation," *Personality and Social Psychology Review,* 11 (2), 167–203.

Cavanaugh, Lisa A., James R. Bettman, and Mary Frances Luce (2015), “Feeling Love and Doing More for Distant Others: Specific Positive Emotions Differentially Affect Prosocial Consumption,” *Journal of Marketing Research*, 52 (5), 657–73.

Chun, HaeEun Helen, Kristin Diehl, and Deborah J. MacInnis (2017), “Savoring an Upcoming Experience Affects Ongoing and Remembered Consumption Enjoyment,” *Journal of Marketing*, 81 (3), 96–110.

Coleman, Nicole Verrochi, Patti Williams, Andrea C. Morales, and Andrew Edward White (2017), “Attention, Attitudes, and Action: When and Why Incidental Fear Increases Consumer Choice,” *Journal of Consumer Research,* 2 (1), 283–312.

Cryder, Cynthia E., Stephen Springer, and Carey K. Morewedge (2012), "Guilty Feelings, Targeted Actions," *Personality and Social Psychology Bulletin*, 38 (5), 607–18.

Dahl, Darren W., Heather Honea, and Rajesh V. Manchanda (2003), “The Nature of Self-Reported Guilt in Consumption Contexts,” *Marketing Letters*, 14 (3), 159–71.

De Hooge, Ilona E., Marcel Zeelenberg, and Seger M. Breugelmans (2007), “Moral Sentiments and Cooperation: Differential Influences of Shame and Guilt,” *Cognition and Emotion*, 21 (5), 1025–42.

Ekman, Paul (1984), “Expression and the Nature of Emotion,” in *Approaches to Emotion*, ed. Klaus R. Scherer and Paul Ekman, Hillsdale, NJ: Erlbaum, 319–43.

Gilbert, Daniel T., Matthew D. Lieberman, Carey K. Morewedge, and Timothy D. Wilson (2004), “The Peculiar Longevity of Things Not So Bad,” *Psychological Science*, 15 (1), 14–19.

Gilovich, Thomas, Victoria Husted Medvec, and Daniel Kahneman (1998), "Varieties of Regret: A Debate and Partial Resolution," *Psychological Review*, 105 (3), 602–05.

Giner-Sorolla, Roger (2001), "Guilty Pleasures and Grim Necessities: Affective Attitudes in Dilemmas of Self-Control," *Journal of Personality and Social Psychology*, 80 (2), 206–21.

Goldsmith, Kelly, Eunice Kim Cho, and Ravi Dhar (2012), "When Guilt Begets Pleasure: The Positive Effect of a Negative Emotion," *Journal of Marketing Research*, 49 (6), 872–81.

Gollwitzer, Peter M. (1999), “Implementation Intentions: Strong Effects of Simple Plans,” *American Psychologist*, 54 (7), 493–503.

Gourville, John T. and Dilip Soman (1998), “Payment Depreciation: The Behavioral Effects of Temporally Separating Payments from Consumption,” *Journal of Consumer Research*, 25 (2), 160–74.

Gross, James J. (1998), “The Emerging Field of Emotion Regulation: An Integrative Review,” *Review of General Psychology*, 2 (3), 271–99.

Gyurak, Anett, James J. Gross, and Amit Etkin (2011), “Explicit and Implicit Emotion Regulation: A Dual-Process Framework,” *Cognition and Emotion*, 3, 400—12.

Izard, Carroll E. (1977), *Human Emotions*, New York: Plenum.

Jager, Gerry, Pascal Schlich, Irene Tijssen, Jiali Yao, Michel Visalli, Cees De Graaf, and Markus Stieger (2014), "Temporal Dominance of Emotions: Measuring Dynamics of Food-Related Emotions During Consumption," *Food Quality and Preference*, 37 (October), 87–99.

Kahneman, Daniel (1995), “Varieties of Counterfactual Thinking,” in *What Might Have Been: The Social Psychology of Counterfactual Thinking*, ed. Neale J. Roese and James M. Olson, Mahwah, NJ: Erlbaum, 375–96.

Kassam, Karim S. and Wendy Berry Mendes (2013), "The Effects of Measuring Emotion: Physiological Reactions to Emotional Situations Depend on Whether Someone is Asking," *PloS One*, 8 (6), e64959.

Kivetz, Ran and Anat Keinan (2006), "Repenting Hyperopia: An Analysis of Self-Control Regrets," *Journal of Consumer Research,* 33 (2), 273–82.

Kivetz, Ran and Itamar Simonson (2002), "Self-Control for the Righteous: Toward a Theory of Precommitment to Indulgence," *Journal of Consumer Research*, 29 (2), 199–217.

Kouchaki, Maryam, Christopher Oveis, and Francesca Gino (2014), "Guilt Enhances the Sense of Control and Drives Risky Judgments," *Journal of Experimental Psychology: General,* 143 (6), 2103–10.

Lascu, Dana N. (1991), "Consumer Guilt: Examining the Potential of a New Marketing Construct," in *Advances in Consumer Research*, Vol. 18, ed. Rebecca Holman and Michael Solomon, Ann Arbor, MI: Association for Consumer Research, 290–95.

Levav, Jonathan and A. Peter McGraw (2009), "Emotional Accounting: How Feelings about Money Influence Consumer Choice," *Journal of Marketing Research*, 46 (1), 66–80.

Lewis, Helen B. (1971), "Shame and Guilt in Neurosis," *Psychoanalytic Review*, 58 (3), 419–38.

Leys, Christophe, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata (2013), "Detecting Outliers: Do Not Use Standard Deviation Around the Mean, Use Absolute Deviation Around the Median," Journal of Experimental Social Psychology, 49 (4), 764–66.

Lieberman, Matthew D., Naomi I. Eisenberger, Molly J. Crockett, Sabrina M. Tom, Jennifer H. Pfeifer, and Baldwin M. Way (2007), "Putting Feelings into Words: Affect Labeling Disrupts Amygdala Activity to Affective Stimuli," *Psychological Science,* 18 (5), 421–28.

Lieberman, Matthew D., Kevin N. Ochsner, Daniel T. Gilbert, and Daniel L. Schacter (2001), "Do Amnesics Exhibit Cognitive Dissonance Reduction? The Role of Explicit Memory and Attention in Attitude Change," *Psychological Science*, 12 (2), 135—40.

Lindsay-Hartz, Janice, Joseph de Rivera, Michael F. Mascolo (1995), "Differentiating Guilt and Shame and their Effects on Motivation," in *Self-Conscious Emotions: The Psychology of Shame, Guilt, Embarrassment, and Pride*, ed. June P. Tangney and Kurt W. Fischer, New York: Guilford Press, 274–300.

Lindsey, Lisa L. Massi (2005), "Anticipated Guilt as Behavioral Motivation: An Examination of Appeals to Help Unknown Others Through Bone Marrow Donation," *Human Communication Research,* 31 (4), 453–81.

Loewenstein, George F., Elke U. Weber, Christopher K. Hsee, and Ned Welch (2001), "Risk as Feelings," *Psychological Bulletin*, 127 (2), 267–86.

Macht, Michael and Dorothee Dettmer (2006), "Everyday Mood and Emotions after Eating a Chocolate Bar or an Apple," *Appetite*, 46 (3), 332–36.

Miceli, Maria and Cristiano Castelfranchi (1998), "How to Silence One's Conscience: Cognitive Defenses Against the Feeling of Guilt," *Journal for the Theory of Social Behavior*, 28 (3), 287–318.

Monteith, Margo J., Leslie Ashburn-Nardo, Corrine I. Voils, and Alexander M. Czopp (2002), "Putting the Brakes on Prejudice: On the Development and Operation of Cues for Control," *Journal of Personality and Social Psychology*, 83 (5), 1029–50.

Morales, Andrea C., Eugenia C. Wu, Gavan J. Fitzsimons (2012), "How Disgust Enhances the Effectiveness of Fear Appeals," *Journal of Marketing Research*, 49 (3), 383–93.

Mukhopadhyay, Anirban and Gita Venkataramani Johar (2009), "Indulgence as Self-Reward for Prior Shopping Restraint: A Justification-Based Mechanism," *Journal of Consumer Psychology*, 19 (3), 334–45.

Neumann, Roland (2000), "The Causal Influences of Attributions on Emotions: A Procedural Priming Approach," *Psychological Science*, 11 (3), 179–82.

Nowlis, Stephen M., Naomi Mandel, and Deborah Brown McCabe (2004), "The Effect of a Delay between Choice and Consumption on Consumption Enjoyment," *Journal of Consumer Research*, 31 (3), 502–10.

Peloza, John, Katherine White, and Jingzhi Shang (2013), "Good and Guilt-Free: The Role of Self-Accountability in Influencing Preferences for Products with Ethical Attributes," *Journal of Marketing*, 77 (1), 104–19.

Peterson, Lloyd and Margaret Jean Peterson (1959), "Short-Term Retention of Individual Verbal Items," *Journal of Experimental Psychology*, 58 (3), 193–98.

Pham, Michel Tuan (2004), "The Logic of Feeling," *Journal of Consumer Psychology*, 14 (4), 360–69.

Pham, Michel Tuan (2013), "The Seven Sins of Consumer Psychology," *Journal of Consumer Psychology*, 23 (4), 411–23.

Prelec, Drazen and George Loewenstein (1998), “The Red and the Black: Mental Accounting of Savings and Debt,” *Marketing Science,* 17 (1), 4–28.

Rachman, Stanley (1993), “Obsessions, Responsibility, and Guilt,” *Behaviour Research and Therapy,* 31 (2), 149–54.

Ramanathan, Suresh and Patti Williams (2007), “Immediate and Delayed Emotional Consequences of Indulgence: The Moderating Influence of Personality Type on Mixed Emotions,” *Journal of Consumer Research,* 34 (2), 212–23.

Read, Daniel and Barbara Van Leeuwen (1998), “Predicting Hunger: The Effects of Appetite and Delay on Choice,” *Organizational Behavior and Human Decision Processes,* 76 (2), 189–205.

Rick, Scott I., Cynthia E. Cryder, and George Loewenstein (2008), “Tightwads and Spendthrifts,” *Journal of Consumer Research,* 34 (6), 767–82.

Robinson, Michael D. and Gerald L. Clore (2002), “Belief and Feeling: Evidence for an Accessibility Model of Emotional Self-Report,” *Psychological Bulletin,* 128 (6), 934–60.

Roseman, Ira J. (1984), “Cognitive Determinants of Emotion: A Structural Theory,” *Review of Personality and Social Psychology,* 5, 11—36.

Sandvik, Ed, Ed Diener, and Randy J. Larsen (1985), “The Opponent Process Theory and Affective Reactions,” *Motivation and Emotion,* 9 (4), 407–18.

Scheier, Michael F. and Charles S. Carver (1977), "Self-Focused Attention and the Experience of Emotion: Attraction, Repulsion, Elation, and Depression." *Journal of Personality and Social Psychology,* 35 (9), 625–36.

Schelling, Thomas (1960), The Strategy of Conflict, Cambridge, MA: Harvard University Press.

Shapiro, Samuel S. and Martin B. Wilk (1965), “An Analysis of Variance Test for Normality (Complete Samples),” *Biometrika,* 52 (3/4), 591–611.

Shiv, Baba and Alexander Fedorikhin (1999), “Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making,” *Journal of Consumer Research,* 26 (3), 278–92.

Smith, Craig A. and Phoebe C. Ellsworth (1985), “Patterns of Cognitive Appraisal in Emotion,” *Journal of Personality and Social Psychology,* 48 (4), 813–38.

Tangney, June Price (1992), “Situational Determinants of Shame and Guilt in Young Adulthood,” *Personality and Social Psychology Bulletin,* 18 (2), 199–206.

Tangney, June Price, Rowland S. Miller, Laura Flicker, and Deborah Hill Barlow (1996), “Are Shame, Guilt, and Embarrassment Distinct Emotions?,” *Journal of Personality and Social Psychology,* 70 (6), 1256–69.

Taylor, Shelley E. (1991), “Asymmetric Effects of Positive and Negative Events: The Mobilization-Minimization Hypothesis,” *Psychological Bulletin,* 110 (1), 67–85.

Thaler, Richard (1985), “Mental Accounting and Consumer Choice,” *Marketing Science,* 4 (3), 199–214.

Tracy, Jessica and Richard W. Robins (2004), “Putting the Self into Self-Conscious Emotions: A Theoretical Model,” *Psychological Inquiry,* 15 (2), 103–25.

VanEpps, Eric M., Julie S. Downs, and George Loewenstein (2016), "Advance Ordering for Healthier Eating? Field Experiments on the Relationship Between the Meal Order–Consumption Time Delay and Meal Content,” *Journal of Marketing Research,* 53 (3), 369–80.

Winkielman, Piotr and Kent C. Berridge (2004), “Unconscious Emotion,” *Current Directions in Psychological Science,* 13 (3), 120–23.

Zemack-Rugar, Yael, James R. Bettman, and Gavan J. Fitzsimons (2007), "The Effects of Nonconsciously Priming Emotion Concepts on Behavior," *Journal of Personality and Social Psychology,* 93 (6), 927–39.

Zemack-Rugar, Yael, Rebecca Rabino, Lisa A. Cavanaugh, and Gavan J. Fitzsimons (2015), "When Donating is Liberating: The Role of Product and Consumer Characteristics in the Appeal of Cause-Related Products," *Journal of Consumer Psychology,* 26 (2), 213–30.

Chapter 2.

THE QUANTITY INTEGRATION EFFECT: WHEN INTEGRATING PURCHASE AND QUANTITY DECISIONS INCREASES SALES

Kristen Elizabeth Duke, On Amir

Rady School of Management, University of California, San Diego, La Jolla, CA, 92093, USA

ABSTRACT

Customers often must decide not only whether to purchase, but also what quantity to buy. The current research introduces and compares the quantity-sequential (QS) selling format, under which shoppers make the purchase and quantity decisions separately, with a quantity-integrated (QI) selling format, under which shoppers simultaneously indicate whether and how much to buy. Although many retailers use the QS format, the QI format often earns higher purchase rates. Across various product offers with over 16,000 experimental observations, this simple change increased average purchase likelihood by 41% and average sales volume by 28%. A field experiment conducted with a large technology company similarly finds a considerable increase in the sales of a major supplies item. The authors detail several psychological mechanisms that contribute to this effect, including a novel mechanism—a "one-click effect"—derived from the sense of closure. In so doing, this work also identifies key moderators of the effect.

INTRODUCTION

Customers often must decide not only whether to make a purchase, but also how many units to buy. Many retailers lay out the purchase process to dictate sequential resolution of these decisions. For example, when ordering pizza online from the largest pizza chain in the world, customers who have selected their desired items are confronted with a pop-up message asking, "Would you like to add one of the following items to your order?" Varied items (e.g., breadsticks, salads, soft drinks) are presented with brief descriptions and prices, each above a "Yes, add to order" button. If customers click on an "add" button, they can later adjust the purchase quantity. Thus, this shop prompts two separate decisions: first, customers choose whether to buy, and then, they choose the purchase quantity. At the outset, making these decisions in separate steps may seem well-guided, as delaying the quantity decision and first choosing only whether to buy could simplify the choice. Accordingly, many leading retailers adopt this sequential strategy (see Appendix for examples). Yet, the current research demonstrates that making a small, seemingly inconsequential change to this process can garner substantially higher sales.

In particular, we compare this quantity-sequential (QS) selling format with a quantity-integrated (QI) selling format, whereby customers answer the purchase and quantity questions simultaneously (see Figure 2.1). To illustrate, consider a slight modification to the pizza website interaction. Rather than displaying only an "Add to order" button, imagine that the shop had displayed buttons of "Add 1 to order," "Add 2 to order," and "Add 3 to order." How would this simultaneous choice impact their propensity to buy? The null prediction, echoed by the widespread industry use of QS selling formats,

is that there should be no difference. Yet, we find that QI selling formats can substantially increase customers' likelihood of making a purchase—even when they buy only one unit. This "quantity integration effect" is simple, robust to various contextual changes, and sizeable, making it a powerful technique for firms to adopt.

We conducted 27 experiments, forming a pooled dataset with over 16,000 observations, to compare the QI and the QS selling formats. A pooled analysis of our dataset estimates a 28% average relative increase in sales volume when changing from a QS to a QI selling format, driven by a 41% average increase in purchase incidence. Thus, this subtle change yields a sizable increase in sales. A large field experiment corroborates the power of quantity integration, finding a substantial lift in conversion and revenue while also demonstrating key boundary conditions. We discuss and assess the role of several underlying psychological mechanisms, including a novel mechanism derived from the sense of closure, that point to additional moderators. In the next section, we review literature related to the importance of framing and choice structure and introduce the candidate mechanisms for this effect.

THEORY

Background

Normative theories of decision-making dictate that individuals' choices ought to be insensitive to the way information is presented (Tversky and Kahneman 1986; von Neumann and Morgenstern 1947). Yet, decades of foundational research find that logically

equivalent but subtly different descriptions of a choice can lead to systematically different decisions (Johnson et al. 1993; Kahneman and Tversky 1979; Puto 1987; Tversky and Kahneman 1981). In the marketing realm, the framing or organization of a purchase process—that is, the "selling format"—can similarly influence what customers buy (Wernerfelt 1994). For example, when customizing a product, consumers who start from a "fully loaded" model and subtract undesired features tend to choose more options than those who start from a base model and add desired features (Park, Jun, and MacInnis 2000). Simply rearranging the order of different attribute decisions can also affect consumers' choices, influencing whether they accept default options (Levav et al. 2010), changing how they search through alternatives (Levav, Reinholtz, and Lin 2012), altering their mental representations of the options (Schrift et al. 2017), and changing the number of items that they purchase (Nowlis, Dhar, and Simonson 2010). Thus, minor changes to the selling format can have important consequences for marketplace decisions.

The current research tests the consequences of changing one feature of the selling format: whether the purchase and quantity decisions resolve in separate steps or simultaneously in one choice. Accordingly, particularly relevant is the idea that separating versus integrating decisions can change how individuals choose. In a key demonstration of this distinction, Dhar and Nowlis (2004) investigate how individuals decide whether to make a purchase (vs. to defer the choice) and which product variant/item to buy. They propose that customers can make these decisions sequentially in a "buy/no-buy mode," by first deciding whether or not to buy (e.g., choosing if they will dine at a restaurant) and then selecting the purchase option (e.g., choosing which specific restaurant), or can reach

these decisions simultaneously in an "unconditional brand-choice mode." Indeed, these distinctions largely mirror the difference between QS and QI. However, Dhar and Nowlis (2004) find that their response modes change what people choose because they shift how people evaluate the options' attributes: the buy/no-buy mode elicits alternative-based evaluations (evaluating each option in isolation), while the unconditional brand-choice mode elicits attribute-based evaluations (comparing attributes across the options; see also Parker and Schrift 2011). The available options in our research are different quantities of the same product, and accordingly they share all the same attributes (unlike the unique product options in this prior work). Thus, mechanisms based on different modes of evaluating attributes are unable to explain the quantity integration effect. Rather, we propose a set of other psychological mechanisms may be at play, which we detail next.

Quantity-Sequential and Quantity-Integrated Selling Formats

We consider two general ways that the QI and QS selling formats can be implemented, each with associated potential mechanisms. In a *theoretical* implementation, nearly all features of the QS and QI selling formats are equated, leaving only their core difference: whether or not customers choose among purchase quantities as part of the first stage of the decision process (see Figure 2.2). A *practical* implementation instead mimics how these selling formats would likely be used in the marketplace (see Figure 2.3). Practical implementations tend to include several additional differences between the QI

and QS selling formats—for example, as reviewed next, the two formats might use different calls to action.

Our experiments employ both theoretical and practical implementations, varying which features differ between formats and which remain constant across formats, to assess the relative impact of each difference in contributing to the quantity integration effect. Next, we discuss potential psychological mechanisms, emphasized in bold text, that stem from the different components of these implementations. We suggest that a combination of these mechanisms likely contribute to the effect. Later in the paper, we evaluate the relative contribution of each mechanism in our set of experiments, although they may contribute differently in other settings.

The call to action. Selling formats must pose a question to customers, either explicitly or implicitly, about the opportunity to purchase. In a practical implementation, retailers using QS might ask, "Would you like to add any drinks to your order?" while those using QI might ask, "How many drinks, if any, would you like to add to your order?" Using different questions may differentially impact behavior because of conversational norms, with customers inferring how they ought to respond (Grice 1975; Hilton 1995; McKenzie and Nelson 2003). This differential language may also signal different social norms (Lieberman, Duke, and Amir 2019). Finally, it may even signal being in different stages of the decision-making process. Any of these processes of **information leakage** from the call to action could contribute to the quantity integration effect.

The response mode. Following the call to action, customers must choose from a set of options, and the number of options tends to differ between the two formats. The QS

format typically carries only two initial options: buying or not buying. In contrast, the QI format offers one non-purchasing option (e.g., "don't buy") but several purchasing options (e.g., "buy 1," "buy 2"). An uncertain decision-maker may **choose randomly** among the choice options (Krosnick 1991); as there are more options involving purchase in a QI (vs. QS) choice, a random chooser would be more apt to purchase in the QI (vs. QS) selling format. Similarly, the greater number of purchase options in QI (vs. QS) could **differentially draw attention** to purchasing over non-purchasing and increase the likelihood that a consumer will choose that option (e.g., Armel, Beaumel, and Rangel 2008).

Having different numbers of choice options in the two selling formats may also prompt extremeness aversion or a **compromise effect** (Simonson 1989; Simonson and Tversky 1992). This could arise if choosing to purchase a small quantity in QI acts as a "compromise option" between not purchasing and purchasing a high quantity, while no such compromise exists in the first stage of a QS choice. Relatedly, customers may draw **choice-relevant inferences** from the choice options presented by the selling format, such as the relative value or magnitude of different purchase quantities (Prelec, Wernerfelt, and Zettelmeyer 1997).

The conceptual difference. The very fact that the QI purchase options include quantities may also activate certain psychological processes that could contribute to the effect. The highest quantity available in QI might imply a **purchasing quantity limit** that could increase customers' likelihood of purchase by signaling the scarcity of the product or the value of the deal (Lessne and Notarantonio 1988; Inman, Peter, and Raghubir 1997) or

by setting an anchor (Wansink, Kent, and Hoch 1998). In contrast, QS does not necessarily offer this quantity limit information.

The two selling formats could also lead to different levels of effort investment. QS customers may choose to **minimize consideration effort** and opt out of the impending second decision of selecting the purchase quantity by choosing not to purchase (Dhar 1996; Shugan 1980). Conversely, the need to respond to two simultaneous questions in QI may lead customers to simplify the decision process via **question substitution** (Kahneman and Frederick 2002; Kahneman 2011), implicitly answering "which quantity do I prefer" rather than whether they truly wish to buy any quantity.

The selling formats may also activate different mindsets. When encountering a QI choice, consumers might adopt a strategy of first considering which quantity they prefer, and only thereafter decide whether to make a purchase. This initial consideration of quantity could induce an **implemental mindset** (Gollwitzer 1999) that then increases propensity to purchase (Xu and Wyer 2007; Dhar, Huber, and Khan 2007). Relatedly, the selling formats may invoke different levels of **construal** (Trope and Liberman 2003). A QI choice option that includes a specific quantity may be more concrete (vs. abstract) than a QS choice option merely indicating purchase. (However, note that a more concrete construal does not necessarily increase likelihood of purchase; in certain circumstances, concrete construals can actually reduce purchase rates: Cho, Khan, and Dhar 2013; Goldsmith, Xu, and Dhar 2016.)

We also propose a novel mechanism that captures the value of settling a decision-making process succinctly and neatly. By definition, QI allows customers to finalize the

choice process in a single action, while QS requires two separate choices, neither of which alone can finalize the purchase process. As a result, the option to purchase in QI (vs. QS) promises a greater feeling of completeness, finality, and resolution—a "one-click effect" . In contrast, we suggest that the action of purchasing feels psychologically incomplete in QS, which reduces customers' willingness to do so. We propose that the underlying psychology of this "one-click effect" shares properties of psychological closure, a satisfying sense of clarity, finality and resolution (Beike and Wirth-Beaumont, 2005; Gu, Botti, and Faro 2013). Prior research demonstrates that lacking closure—feeling as if things are "psychological unfinished and unresolved" (Beike et al. 2009)—is aversive and uncomfortable (Beike, Kleinknecht, and Wirth-Beaumont 2004), and accordingly individuals often actively seek out closure when deciding how to act (e.g., Webster and Kruglanski 1994). In the marketing domain, seeking closure drives consumers to make final purchase decisions rather than continue examining the available options (Kupor, Reich, and Shiv 2015), and to respond more favorably to immediate (vs. delayed) marketing promotions, such as instant discounts versus mail-in rebates (Kim 2013). We suggest that being able to finish the choice process in "one click" lends a sense of closure and finality to QI that increases customers' willingness to purchase. As such, we test this account using measures and manipulations of the sense of closure as a proxy for this notion of a "one-click effect."

DATA OVERVIEW

First, we present two sample incentive-compatible lab experiments: one using a typical theoretical implementation of the selling formats (1A), and another in which customers do not answer explicit purchase and quantity questions, but rather indirectly answer them by choosing how to navigate through a website (1B). Second, we highlight experiments testing our newly proposed theoretical mechanism of closure (2A—2C). Thereafter, we present a pooled analysis of 26 experiments, which estimates the magnitude of this quantity integration effect and identifies moderators of it. Following, we present a field experiment demonstrating the effect and delineating a boundary condition. We conclude by evaluating the evidence for and against various psychological mechanisms and discussing implications for theory and practice.

EXPERIMENT 1A: SAMPLE LAB EXPERIMENT, TIME-CONTROLLED

Experiment 1A demonstrates the quantity integration effect in an incentive-compatible lab study. It employs a theoretical implementation, holding constant all differences between the formats aside from their conceptual difference. Additionally, participants must spend equal amounts of time making choices in both formats, regardless of what they choose. This time control was designed to test two candidate mechanisms derived from differing levels of effort investment: QS prompting consideration effort minimization, and QI prompting question substitution. If participants are forced to spend time on the decision, these mechanisms should have less of an opportunity to operate. The persistence of the effect in this situation casts doubt on these mechanisms.

Method

Four hundred three Mechanical Turk workers (46% female; M_{age} = 35.6 years, SD_{age} =11.5 years) were randomly assigned to one of two conditions (selling format: QS vs. QI) in a between-subjects design. All participants read that they would make a purchase decision and that one worker would be selected to receive a monetary bonus and any products that he/she purchased. They then learned that the main sections of the survey "will take a total of 30 seconds, regardless of what you choose. So, please take your time and choose carefully."

Next, participants read, "If you are selected, you will receive a $20 bonus. If you'd like, you can use this money to purchase Lemome notebooks (pictured below) for $6 per notebook (current price on Amazon: $9). You can buy up to 3 notebooks. Any money you do not spend on notebooks will be given to you as a bonus." (See Appendix for materials.) All participants then answered, "What would you like to do?" QS participants chose between "Not buy any notebooks" and "Buy," while QI participants chose among, "Not buy any notebooks," "Buy 1 notebook," "Buy 2 notebooks," and "Buy 3 notebooks." There was a hold timer on this page requiring participants to spend at least 10 seconds on it.

On the following page, QS participants who had chosen to purchase then selected the quantity, answering, "You indicated you will make a purchase. What would you like to do?" with choice options matching the QI condition: "Not buy any notebooks," "Buy 1

notebook," "Buy 2 notebooks," and "Buy 3 notebooks".[3] On the same page, all participants (in both formats, regardless of whether they purchased) also answered a free-response question, "Why did you make the purchase decision you made?" There was a hold timer on this page that required participants to spend at least 20 seconds on it. Thus, regardless of what decisions participants had made and regardless of their condition, they had to spend a required minimum amount of time. Finally, participants provided demographic information.

Results and Discussion

QS participants were significantly less likely to purchase (29.40%) than were QI participants (39.60%; $\chi^2(1, N = 403) = 4.69$, $p = .03$, $\varphi = .11$); quantity integration promoted a 35% relative increase in purchase likelihood. Among purchasers, QS participants bought a similar number of notebooks (M = 1.41, SD = .75) as did QI participants (M = 1.31, SD = .59; t(137) = -.83, $p = .41$), suggesting that the selling format did not discernibly affect the purchase quantity. Overall, the QS condition sold fewer total notebooks (83; N = 201) than did the QI conditions (105; N = 202). In sum, changing from a QS to a QI format increased total volume sold by 27%.

EXPERIMENT 1B: SAMPLE LAB EXPERIMENT, WEBSITE NAVIGATION

[3] Note that one participant chose to buy in the first choice but changed her mind in the second; coding her response as either a purchase or non-purchase does not affect the results.

In many retail contexts, quantity integration can be implemented by manipulating the call to action question, as in Experiment 1A. Yet, quantity integration may also arise when customers navigate a website and choose to add items to their shopping carts. There, customers do not answer explicit questions, and do not choose "No" if they do not wish to buy. Rather, those navigating in a QS environment may decide whether to add (a default of) one item to their carts, and later may adjust this quantity. If they do not wish to purchase, they can simply navigate to a different product or different area of the website. In this context, quantity integration can be facilitated by simply replacing an "add to cart" option with a set of options specifying quantities (e.g., "add 1 to cart" and "add 2 to cart"). Experiment 1B tests the effects of quantity integration in this paradigm by guiding participants through a scenario of shopping on the website of a large, same-day grocery delivery service: Instacart (https://www.instacart.com).

Method

Four hundred thirty-three undergraduates (59% female; M_{age} = 21.1 years, SD_{age} = 2.5 years) from a large west coast university participated in a laboratory study for course credit. Participants were randomly assigned to one of two conditions (selling format: QS vs. QI) in a between-subjects design. All participants navigated through copies of the Instacart website and arrived at the produce section. They then simulated their next website action. Specifically, they read, "You hover over the avocado item, and see the following options pop up. Please imagine you are actively navigating this website, and click the next

thing that you would click in this situation." This text was displayed above a modified screenshot that varied by condition and reflected the interface logic of the actual website.

Participants in the QS condition saw the screenshot displayed in the left panel of Figure 2.4. They could click "Fresh Vegetables" (which on Instacart's website, brings the patron to the top of the produce section's landing page), the heart icon (which adds the item to a list of "favorites"), the arrow button (which scrolls to show additional items in the product category), or "+Add" (which adds one unit of the item to the shopping cart)[4]. For simplicity in our experiment, this initial click was the only decision participants made; there was no second choice allowing participants to adjust the quantity (and accordingly, we will only analyze purchase likelihood). Participants in the QI condition saw a similar screenshot (Figure 2.4, right panel). The only difference was that that "+Add" button was replaced with 3 options: "+Add 1," "+Add 2," and "+Add 3." In this way, in a single click, participants could both specify their purchase intent and their intended quantity.

One possible concern with this QI modification is that it could make the choice or the website more complicated or less user-friendly, reducing customers' willingness to visit at all. To address this possibility, all participants next answered, "How likely would you be to use such a service?" (0: Not at all to 100: Extremely likely). Finally, they indicated if they have ever used Instacart or a similar service and provided demographic information.

[4] Note that in this QS selling format, a customer cannot specify the desired quantity in a single click (unlike the QI format). A customer would need to click the add button repeatedly, corresponding to his/her desired quantity.

Results and Discussion

Clicks. QS participants were significantly less likely to add the item to their cart (30.88%) than were QI participants (43.06%; $\chi^2(1, N = 433) = 6.89$, $p = .009$, $\varphi = .13$), a 39% relative increase. This difference seemed to arise from participants trading off between adding the item to the cart and clicking the "arrow" button: QS participants were significantly more likely to click the arrow (62.2%) than were QI participants (49.6%; $\chi^2(1, N = 433) = 6.89$, $p = .009$, $\varphi = .13$). Participants were equally likely to engage with the page in other ways, clicking the heart button at equal rates (QS: 5.5% vs. QI: 4.6%; $\chi^2(1, N = 433) = .18$, $p = .67$, $\varphi = .02$) and clicking the product category name at equal rates (QS: 1.4% vs. QI: 2.8%; $\chi^2(1, N = 433) = 1.04$, $p = .31$, $\varphi = .05$).

Usage Likelihood. This modification did not appear to alienate potential customers, as participants reported they would be equally likely to use the site in both formats (QS: M = 37.79, SD = 28.69 vs. QI: M = 36.10, SD = 29.37; $t(431) = -.60$, $p = .55$). A regression controlling for whether or not participants added the item to their cart reveals the same pattern, with no significant effect of selling format ($\beta = 2.14$, $t(430) = .76$, $p = .48$), nor of whether participants added the focal item ($\beta = 3.74$, $t(430) = 1.28$, $p = .20$). An additional regression revealed no interaction between these two factors ($\beta = -.31$, $t(429) = -.05$, $p = .96$). Thus, the QI format increased participants' likelihood of adding an item to their cart without affecting their willingness to use the service itself.

EXPERIMENTS 2A-2C: SUPPORT FOR A NOVEL MECHANISM, CLOSURE

Experiments 2A-2C use theoretical implementations of the selling formats to test the novel proposed mechanism: differences in the sense of closure associated with a "one-click effect." Below, we present a summary of each experiment; additional details can be found in the Appendix.

Experiment 2A: Reported Sense of Closure After Choice

This experiment tests whether the formats indeed differ in the sense of closure they provide after the initial purchase choice. We predicted that QS customers would experience less closure after choosing to "buy" than QI customers would feel after choosing to "buy N units."

Method. Eight hundred Mechanical Turk workers (50% female; M_{age} = 35.7 years, SD_{age} =11.8 years) were randomly assigned to either the QS or QI condition. All participants read that one worker would be selected to receive a bonus and any products purchased. The purchase option was up to 4 bags of Lindt milk chocolate truffles. All participants answered, "What would you like to do?" QS participants chose between "Not buy any truffles" and "Buy truffles," while QI participants chose among, "Not buy any truffles," "Buy 1 bag of truffles," "Buy 2 bags of truffles," "Buy 3 bags of truffles," and "Buy 4 bags of truffles." After this initial decision, all participants responded to three measures capturing their sense of closure with their choices (adapted from Gu, Botti, and

Faro 2013; $\alpha = .74$). On the next screen, QS participants who chose to purchase selected a quantity using the same quantity options as QI participants.

Result: Purchase. QS participants were significantly less likely to purchase truffles (26.57%) than were QI participants (35.91%; $\chi^2(1, N = 800) = 8.13$, $p = .004$, $\varphi = .10$). Among those who purchased, QS participants purchased a similar number of bags of truffles (M = 1.65, SD = 1.00) as did QI participants (M = 1.61, SD = .89; t(248) = -.33, $p = .74$). In sum, the QS condition sold fewer total bags of truffles (175; N = 399) than did the QI condition (232; N = 401); in this context, quantity integration increased the total sales volume by 33%.

Result: Sense of Closure. Unsurprisingly, purchasers reported experiencing less closure than non-purchasers, who have completely ended the decision process ($\beta = -19.78$, $t(797) = -12.93$, $p < .001$). Furthermore, QS participants felt less closure than QI participants did ($\beta = -4.60$, $t(797) = -3.24$, $p = .001$). Importantly, the difference in closure across formats was particularly pronounced among participants who had made a purchase (*purchasers:* QI: M = 70.78, SD = 22.16 vs. QS: M = 62.71, SD = 22.18; t(248) = 2.84, $p = .005$; *nonpurchasers:* QI: M = 88.26, SD = 18.61 vs. QS: M = 85.21, SD = 19.04; t(548) = 1.90, $p = .06$; interaction: $\beta = -1.25$, t(796) = -1.63, $p = .10$). Thus, purchasing in the QI format provided greater closure than did purchasing in the QS format, consistent with our theorizing.

Experiment 2B: Moderation by the Need for Cognitive Closure

We propose that the asymmetry in closure is a key driver of the quantity integration effect. Accordingly, this experiment tests for heterogeneous treatment effects based on individuals' sensitivity to closure. If the closure afforded by QI contributes to the effect, it should increase purchase incidence the most among individuals who tend to seek closure when deciding how to act (i.e., those high in the need for cognitive closure; Webster and Kruglanski 1994). In contrast, individuals who are minimally influenced by the allure of closure should show an attenuated quantity integration effect.

Method. Six hundred three Mechanical Turk workers (51% female; .5% other; M_{age} = 35.2 years, SD_{age} = 10.9 years) were randomly assigned to either the QS or QI condition. All participants imagined ordering pizza delivery online and encountering an advertisement to add Coca-Cola to their order. All participants answered, "Please indicate what you would do in this situation." For QS participants, the options were "Not buy any" and "Buy." For QI participants, the options were "Not buy any," "Buy 1 bottle," "Buy 2 bottles," and "Buy 3 bottles." Thereafter, QS purchasers selected a quantity using the same options as QI participants. All participants then completed the 15-item abridged Need for Closure Scale (NFCC; $\alpha = .88$; Roets and Van Hiel 2011; there was no effect of condition on NFCC scores; $\beta = .07$, SE = .06, $t(601) = 1.27$, $p = .21$).

Results: Purchase. QS participants were again significantly less likely to purchase any soda (22.19%) than were QI participants (30.56%; $\chi^2(1, N = 603) = 5.45$, $p = .02$, $\varphi = .10$). Among those who chose to buy, QS and QI participants chose similar quantities (QS: M = 1.27, SD = .59; QI: M = 1.22, SD = .51; $t(157) = -.58$, $p = .56$). All told, the QS

condition yielded "sales" of fewer total bottles of soda (85; N = 302) than did the QI condition (112; N = 301).

Results: Moderation by NFCC. A binary logistic regression predicting purchase from the selling format (QS vs. QI, effect-coded), NFCC (mean-centered), and their interaction, with Coca Cola purchasing frequency as a covariate, yielded a significant NFCC x selling format interaction (β = -.29, SE = .14, z = -2.17, p = .03). A floodlight analysis (Spiller et al. 2013) revealed that there was a significant negative effect of being in the QS (vs. QI) condition on purchase for any NFCC score greater than 3.80 (out of 6; 60.03% of the data; β = -.37, t = -1.96, p = .05), but no effect for any value of NFCC less than 3.80 (39.97% of the data). Thus, quantity integration increased purchase intent for participants who are most likely to seek closure (the majority of individuals), but had no effect among participants low in this tendency.

Experiment 2C: Moderation by Inducing the Sense of Closure

This experiment tests whether inducing the sense of having closure—or conversely, the sense of lacking it—can moderate the quantity integration effect. Instilling a feeling of lacking closure, which should drive individuals to seek closure in their choices, should yield a large quantity integration effect. In contrast, offering a sense of closure prior to deciding should weaken this drive and attenuate the effect.

Method. Six hundred fifty-five undergraduates (59% female; M_{age} = 21.2 years, SD_{age} = 2.9 years) were randomly assigned to one of four conditions in a 2 (selling format:

QS vs. QI) X 2 (prime: lack of closure vs. high sense of closure) between-subjects design. Participants completed a two-minute writing task designed to prime either a sense of lacking closure or a high sense of closure (see full text and pretest in the Appendix). Next, they learned they would receive $1 in quarters to keep or to use for buying gum at 50 cents/pack. All participants answered, "What would you like to do?" For QS participants, the options were "Make a purchase" and "Not make a purchase." For QI participants, the options were "Purchase 1 pack," "Purchase 2 packs," and "Not make a purchase." Thereafter, QS purchasers selected the quantity.

Results: Purchase. Overall, QS participants were significantly less likely to purchase any packs of gum (42.38%) than were QI participants (53.52%; $\chi^2(1, N = 655) = 8.14$, $p = .004$, $\varphi = .11$). Purchasers bought similar quantities in both formats (QS: M = 1.71, SD = .46; QI: M = 1.65, SD = .48; t(310) = -1.06, $p = .29$). Overall, the QS condition sold fewer total packs of gum (234; N = 328) than did the QI condition (289; N = 327): changing from a QS to a QI format increased total sales by 25%.

Results: Moderation. A binary logistic regression revealed a main effect of selling format ($\beta = -.24$, SE = .08, z = -2.97, $p = .003$), whereby QI (vs. QS) participants were more likely to buy, and a main effect of the prime ($\beta = -.17$, SE = .08, z = -2.15, $p = .03$), whereby participants in the lacking closure condition were less likely to buy. Importantly, these effects were qualified by a significant prime x selling format interaction ($\beta = -.17$, SE = .08, z = -2.16, $p = .03$). Within the lacking closure conditions, participants in the QS format were significantly less likely to make a purchase than those in the QI format ($\beta = -.82$, SE = .23, z = -3.57, $p < .001$). However, this effect was attenuated in the high sense of

closure conditions (β = -.13, SE = .22, z = -.58, p = .56). Furthermore, this closure manipulation selectively affected the QS format: priming a lack of closure significantly decreased purchase rates in the QS format (β = -.69, SE = .23, z = -3.01, p = .003), but this prime had no effect in the QI format (β = .001, SE = .22, z = .005, p > .99), consistent with our theorizing.

Together, these experiments support the role of closure in the quantity integration effect. Next, we test for other key drivers by conducting a pooled analysis of all experiments.

POOLED ANALYSIS

In total, we conducted 27 experiments[5], with a total of 16,431 observations, varying different product offers, price points, and contextual features to demonstrate the robustness of the proposed quantity integration effect and identify moderators of it. Thus, within an experiment, we often varied an additional factor beyond the selling format manipulation (e.g., varying whether participants were choosing for themselves vs. for another individual). Because each factor is orthogonal to the selling format manipulation, we split each experiment that follows a two-by-two design, treating it as two separate experiments (McShane and Böckenholt 2017). Thus, in the pooled analysis, each "experiment" represents a simple two-condition design (selling format: QS vs. QI). The resultant data include 52 "experiments" testing the quantity integration effect. We present a

[5] Note that this count excludes Experiment 1B, which followed a slightly different design and is therefore not included in this analysis.

complementary single-paper meta-analysis (McShane and Böckenholt 2017) in the Appendix that yields results comparable to this analysis.

Experimental Details

Supplemental Table 1 includes methods details for all experiments, including those presented in the main paper. All experiments compare customer purchase rates in the QI and QS selling formats. Participants made purchase decisions about a variety of products: raffle tickets, soda, candles, soaps, pens, chocolates, and more. While most products were mundane and inexpensive (from 5 cents to $9.99 per item), some were costly (e.g., $200). In addition, participants made these decisions in a variety of contexts, such as when the exact product was randomly chosen, or when the choice was on behalf of another individual.

Results: The Quantity Integration Effect

Figure 2.5 presents the average difference in purchase rates in each split experiment. In most experiments, the QI selling format yielded significantly higher rates of purchase than did the QS format, ranging from a -1% to a 33% percentage-point difference. Across all data points, changing from a QS to a QI format yielded a 41% relative increase in purchase likelihood (changing from 36.91% in QS to 51.95% in QI).

Table 1 presents logistic mixed-effects models and linear probability mixed-effects models predicting purchase. We include split-experiment random effects to account for variation attributable to unique experiment and/or product characteristics (e.g., time of day, product category) and individual random effects to account for repeated measures on certain participants (more details on this are provided next). These analyses reveal that the QI selling format resulted in significantly higher purchase incidence than the QS selling format, even when controlling for: whether the maximum purchase quantity was explicitly mentioned, the maximum quantity itself, whether the same question was used in both conditions, whether the purchase was standalone or part of an existing purchase (e.g., a suggested "add-on" item to an already-started order), and the product's price. These analyses also identify overall drivers of purchase, across formats. Consistent with prior research, participants were generally more likely to buy when a maximum quantity limit was (vs. was not) explicitly mentioned (e.g., Wansink et al. 1998). Unsurprisingly, individuals were also more likely to purchase lower-priced (vs. more expensive) items.

Moderators. As can be seen in Figure 2.5, there is substantial heterogeneity in the size of this effect. Accordingly, Table 2.1 also reveals moderators of the effect. Specifically, the quantity integration effect was smaller in the experiments where the exact same question (call to action) was used in both selling formats (vs. when the formats asked different questions), although note that the effect was still quite significant ($p < .001$) in experiments using the same question in both formats. Second, the effect was also larger among more expensive products. We elaborate on what these moderators reveal about the psychological drivers of the effect in Section 9. Note that explicitly mentioning the

maximum purchase quantity did not moderate the effect, nor did the value of that maximum quantity. Whether the purchase opportunity was presented as part of an existing purchase also did not influence the size of the effect.

Within-person effects. Some individuals participated in multiple different experiments, allowing for an examination of within-person effects. We had 2,999 individuals who participated in two or more experiments[6], contributing a total of 9,401 observations. A binary logistic regression predicting purchase among these participants, with study fixed effects, revealed a significant positive effect of the QI format ($\beta = .73$, $z = 16.17$, $p < .001$).[7] Similarly, 1,397 participants contributed at least 3 observations (for a total 6,067 observations); this sample also showed a significant positive effect of the QI format ($\beta = .74$, $z = 13.25$, $p < .001$). Even the nineteen individuals who supplied 10 or more observations (for a total of 210 observations) showed this effect ($\beta = 1.44$, $z = 3.80$, $p < .001$); this is true at any level of multiple observations up to 10 (i.e., including participants with at least 4 observations; with at least 5 observations; etc.; *p*'s < .001). Thus, the same exact person was substantially more likely to buy in the QI format than the QS format, repeatedly. This lack of "wearing off" across repeated encounters with these selling formats speaks to the robustness of this effect.

[6] Note that one experiment required participants to make choices for multiple products (Studies K1-K4). For simplicity, these observations were excluded from analyses, although including them does not change the results.

[7] Including random effects for participant reveals a consistent effect of format, $\beta = -.92$, $z = -16.33$, $p < .001$. A model instead including fixed effects for participant fails to converge, but including fixed effects in a linear probability model yields convergent results, $\beta = -.16$, $t(6241) = -15.27$, $p < .001$. These patterns persist for any subset of repeat participants (i.e., those with at least 3 observations; those with at least 4 observations, etc.). Note also that the effect persists, of course, among only participants who participated in just a single study, $\beta = .68$, $z = 20.14$, $p < .001$.

Results: Total Quantity Purchased

Overall, conditional on purchase, participants purchasing in the QI format bought significantly fewer units on average than did those purchasing in the QS format (with no controls: $\beta = -.23$, SE = .04, $t(7268) = 5.71$, $p < .001$; with Product fixed effects: $\beta = -.21$, SE = .03, $t(7254) = 6.47$, $p < .001$; with Study fixed effects: $\beta = -.20$, SE = .03, $t(217) = 6.25$, $p < .001$). Our theory does not offer a particular prediction about quantity purchased, but this may represent a selection effect: if the QS format adds a slight force against making a purchase, it is most likely to discourage the customers who are closest to indifference—those with a weaker need or preference, who are likely to buy the lowest quantity.

Importantly, the QI (vs. QS) format still yielded a higher *average* purchase quantity (i.e., including non-purchasers who buy 0 units; with no controls: $\beta = .24$, SE = .02, $t(16429) = 9.85$, $p < .001$; with Product fixed effects: $\beta = .24$, SE = .02, $t(16415) = 10.85$, $p < .001$; with Study fixed effects: $\beta = .24$, SE = .02, $t(16378) = 11.08$, $p < .001$). In terms of total sales volume, changing from a QS to a QI format yielded a 28% average increase in total units sold (from 0.87 to 1.12 units per customer).

FIELD EXPERIMENT

Having demonstrated the robustness of this effect in the lab, we partnered with a large technology company to test quantity integration and its boundaries in the field.

Field Experiment Setup

The company allowed us to suggest an alternative interface to a key page on their website. This interface had been selected as the optimal interface after A/B testing and had remained the strongest performer for two years thereafter. It uses a QS selling format where the purchase and quantity decisions are resolved in separate steps/with separate buttons. Note, however, that in this case, the two steps are displayed right beside one another. Customers specify the desired quantity using "+" and "–" buttons and click an adjacent "Add to Cart" button to initiate the purchase process (see Figure 2.6). This presents a more conservative point of comparison for a QI change: can quantity integration increase sales, even in a case where the two decisions in the QS format are already presented quite closely to one another? Our proposed QI implementation combined the quantity and purchase decisions into a single choice. As in our experiments, it presented customers with multiple purchase options, each specifying a different quantity (see Figure 2.7). This context also serves as a more conservative test because it presents many product offerings on the same page, potentially diluting the effect of the selling format.

Two types of products were included in the test. Per an interview with company personnel, Product A, our focal test product, tends to be purchased by business customers who typically carry a stock of this product. When approaching the website, then, potential

Product A purchasers may be open to the idea of buying, but are not necessarily committed to it. In contrast, company personnel stated that Product B tends to be purchased by individual consumers at the moment of need—after their supply has been depleted. These shoppers have, effectively, proceeded past the initial stage in a QS decision, and the only remaining question is which quantity is appropriate. Product B is also a category with multiple variants that are usually purchased as complements: customers typically need to purchase two different types of Product B (but this is not the case for Product A). As such, even in the QI format, selecting to purchase one variant of Product B may not provide a sense of closure, as most customers need to continue with the shopping process to buy the second variant. For these reasons, quantity integration had the potential to lift sales of Product A, but should have little impact on Product B.

On March 6, 2019, the company launched an A/B test on their website comparing the two interface presentations. This test was in place for all website traffic and ran for 2.5 weeks, ending on March 23, 2019[8]. In the data, we observe visits, orders, the type of items ordered, revenues, and units purchased. Thus, for our key metric of purchase likelihood (conversion rates), we follow the company's guidelines and compare purchases to visits that did not result in purchases. Note that this comprises all visitors to the website, including those who never searched for the focal products being tested and were not exposed to the key manipulation, and accordingly represents an intent-to-treat sample.

[8] On the heels of this initial test's success, the company decided to conduct an additional test during two follow-up periods that spanned two product sales. Given that this secondary test was exploratory and not planned *a priori*, we discuss its results and a possible interpretation of them in the General Discussion.

Field Experiment Results

Conversion. We first examine sales of Product A, our focal product. When changing from the QS to the QI format, conversion rates of Product A increased 12.70%, a statistically and economically significant increase ($\chi^2(1) = 8.36$, $p = .004$)[9].

Revenue. Changing to QI promoted a 15.68% increase in revenue per visit (RPV) for Product A. This was driven both by the increase in purchase incidence and a 2.06% increase in units per order.

Product B. As expected, quantity integration had little effect on Product B. Conversion rates non-significantly decreased 2.13% when changing from the QS to the QI format ($\chi^2(1) = 1.61$, $p = .20$). Accordingly, RPV fell 1.02%; QI also saw a 0.58% increase in the average number of units per order.

Field Experiment Discussion

Quantity integration substantially increased sales and revenue in the field. In response to this result, the company is changing this page for the focal test product from its existing QS format to the new QI format.

At the same time, quantity integration had no effect on the sales and revenue of a product for which (a) customers shop only when they truly need it, and (b) customers typically need to purchase complementary variants at the same time. We take these results

[9] We omit the reporting of the actual Ns and individual conversion rates at the request of the company.

to mean that quantity integration can be a powerful tool for promoting purchase, but it may be less effective in cases where customers have already decided to buy or where choosing to buy requires the purchase of multiple complements, which negates the power of decision resolution in a single step. In the next section, we use evidence from this field experiment and from our lab experiments to evaluate the psychological drivers of the effect.

ASSESSING THE ROLE OF DIFFERENT FEATURES AND MECHANISMS

Here, we draw on our empirical evidence to evaluate how various differences between QS and QI may contribute to the quantity integration effect in the settings we test. We focus on the key aspects of selling format implementation: the call to action (how customers are invited to buy), the response mode (the number and arrangement of choice options), and the conceptual difference (whether or not quantities are involved in the customer's purchase decision). It is important to note that we can only evaluate the operation of these mechanisms in the specific contexts, implementations, and methodologies we tested. We hold that even mechanisms that do not seem to contribute to the effect in our studies may contribute to it in other settings, and view these possibilities as directions for further research.

The Call to Action

Using different calls to action in the two selling formats, as in a practical implementation, can influence behavior by prompting inference-making. Indeed, the pooled analysis revealed that using different questions in the formats yielded a larger quantity integration effect, lending credence to this mechanism. However, the effect still arose when using theoretical QI and QS implementations with identical calls to action. Further, the effect arose even when no explicit questions were asked (lab Experiment 1B and field experiment). Thus, inference-making (about conversational norms, social norms, or the perceived stage in the decision-making process) seems to contribute to the effect, but is not the primary driver.

The Response Mode

Random choice. QI (vs. QS) could mechanically earn higher purchase rates if participants choose randomly, as QI tends to have more options involving purchase. However, the quantity integration effect arose in incentive-compatible lab experiments and a field experiment where customers spent real money to make purchases, situations where they are unlikely to choose randomly. Moreover, the pooled analysis compared the magnitude of the effect across studies with different maximum purchase quantities. When there is a higher maximum purchase quantity, there are also necessarily more QI purchasing options, so random choice would predict a larger quantity integration effect when the maximum quantity is higher. Yet, there was no relation between purchase rates and the maximum quantity, casting doubt on this account.

Differential attention. The greater number of purchase options in QI (vs. QS) could also draw greater attention, and accordingly increase their selection, but this account would also predict a larger quantity integration effect with a higher maximum purchase quantity. Further, this account would also suggest that the visual presentation of the option should matter. However, participants were equally likely to purchase when QI was presented in the drop-down format—where the different purchase options are initially hidden—as when QI was presented in the sliding scale format—which makes the purchase options more salient (Appendix Experiment V). This mechanism would also predict an effect for Product B in the field experiment, but there was no evidence for one. Together, these results deem this explanation unlikely.

Compromise effect and Choice-relevant inferences. If QI respondents were exhibiting a compromise effect, we would expect different average purchase quantities between the two formats. However, in many experiments, we observed a significant shift in purchase likelihood without any change in the quantity purchased between formats. Further, Appendix Experiment R shows that the effect can arise even in a situation where the maximum purchase quantity is one unit, where there is no option that could serve as the compromise. Prior research also shows that the compromise effect arises in part because the choice set leaks information about the options themselves (Prelec, Wernerfelt, and Zettelmeyer 1997), a process that may have some positive contribution to the quantity integration effect. However, we observed a difference in the effect of quantity integration for the two products in the field experiment, despite implementing the formats identically

for both products. Coupled with the above evidence, such a process is unlikely explain the effect.

The Conceptual Difference: When Quantities Are (Vs. Are Not) Involved

Quantity limit information. QS may not necessarily make participants aware of any quantity limit, while QI necessarily does. To test whether this influenced behavior, we conducted a set experiments where QS did not explicitly mention a maximum quantity, and others where it did. The pooled analysis found no difference between such cases, suggesting that mechanisms tied to this feature (anchoring on quantities or signaling of scarcity or value) cannot explain the quantity integration effect.

Differential effort investment. We considered the possibility that QS customers might minimize consideration effort and opt out of the (less certain) second quantity decision by choosing not to buy at all. To test this, we conducted an experiment where we put the purchase and quantity decisions immediately after one another on the same page, so participants could see all of the specifics of the purchase process. This manipulation had no effect on the magnitude of the quantity integration effect (Appendix Experiment U). Similarly, the field experiment's QS format had the two decisions directly beside one another and still yielded a large effect. Further, Experiment 1A demonstrated that even when task times are strictly controlled, the effect still arises. Finally, difference in effort levels cannot explain the difference between the two products in the field experiment. Thus, it seems unlikely that differences in effort could drive this effect.

Question substitution. It is possible, particularly in hypothetical experiments, that participants could perceive that they ought to be answering different questions in the two formats (i.e., QI participants thinking they should answer "which quantity is best" rather than "what would I like to buy"). However, this is an unlikely explanation in consequential choice experiments, where participants' decisions were carried out, and especially in a field experiment, where customers did not know they were participating in a study and there was no "question" prompt.

Implemental and deliberative mindsets or Level of construal. Appendix experiment W was designed to test both of these mechanisms. It finds that inducing an implemental mindset—which should attenuate the effect under this mindset account—does not moderate the effect (and if anything, it slightly accentuated it), so this explanation is unlikely. Similarly, the same experiment finds no support for construal level differences.

Sense of closure. As shown in Experiments 2A-2C, the sense of closure seems to play an important role. Purchasing offers greater closure in the QI format than in the QS format, which enhances the appeal of buying, accordingly increasing purchase.

GENERAL DISCUSSION

In countless marketplace contexts, customers must decide both whether and how much to buy. They can reach these decisions either sequentially or simultaneously in a single choice. Importantly, marketers can easily change how customers approach these decisions by slightly modifying the selling format. We compare the quantity-sequential

(QS) selling format, that separates the purchase and quantity decisions, to the quantity-integrated (QI) selling format that combines the two. Relative to QS, QI substantially increases purchase likelihood, an effect that appears to be multiply determined by various mechanisms derived from differences in the call to action, response mode, and conceptual structure of the two formats. Next, we discuss the practical and theoretical implications of these findings.

Practical Considerations and Possible Extensions

Most simply, our results suggest that marketers should pay careful attention to how they invite customers to buy. In the digital realm, online retailers have tools at their disposal to encourage simultaneous purchase and quantity considerations. Merchants may simply change their purchase appeals to provide QI choice options (e.g., "Add 1 cookie," "Add 2 cookies") rather than QS ones (e.g., "Add to cart"). Our field experiment demonstrates that changing to a QI format, even from a QS format that already had the purchase and quantity decisions presented beside one another, can have a substantial impact on real sales.

The focus of this work was in identifying how the structure of a selling format can change what customers choose. However, it is possible that other marketer interactions may similarly shape how customers think about the purchase decision and accordingly change their willingness to buy. For example, marketers may need to be mindful of the types of questions they ask customers—in emails, in in-person interactions, in

advertisements and more—and whether they encourage a QI- or a QS-oriented decision process. Consider the difference between an ad appeal that is framed to encourage a QS decision process (e.g., "Get it while it's hot!") and one framed to encourage a QI decision process (e.g., "Grab a few – one or two?"). We posit that appeals encouraging QI considerations may have the potential to increase purchase rates, an interesting question for future research.

In the current research, we tested products that may plausibly be purchased in multiple units (e.g., bottles of soda, packs of gum). However, we posit that attempts to encourage purchase via quantity integration may backfire, eliciting reactance, if they are perceived as attempts at persuasion (Clee and Wicklund 1980; Fitzsimons and Lehmann 2004). In a case where it is quite unusual for people to purchase multiple quantities (e.g., when buying a new car or smartphone), consumers may perceive QI appeals as strange or malicious. Thus, we caution that the quantity integration effect may be limited to situations where a customer could reasonably expect to buy more than one unit. However, we still found a strong quantity integration effect even in situations where, in reality, nearly all shoppers ended up purchasing only one unit (e.g., one bottle of Coke when ordering a pizza online). Thus, we propose it is the mere reasonableness of asking about multiple quantities, rather than whether a customer actually intends to buy multiple units, that qualifies the applicability of this effect.

Our experiments also primarily tested situations where customers do not, at the outset, already have the explicit goal to purchase the item in question. We did this for an important reason: if customers already have the goal to purchase a product, they may be

past the first decision in a QS process (deciding to buy). This was underscored by the field experiment, where the product that customers tend to only buy when they urgently need it showed no quantity integration effect (although note that we cannot tease apart this boundary from the alternative possibility that the need to purchase multiple product variants limited this effect). We suggest that quantity integration is most applicable to unplanned purchases, which make up approximately half of what shoppers buy (Kollat and Willett 1967; Stilley, Inman, and Wakefield 2010). However, even in cases with partial (but incomplete) intentions—such as a goal to make a purchase in a particular product category, with no resolution about the specific brand or SKU—changing the selling format could potentially increase the ease and frequency with which customers make a purchase (vs. defer the choice or navigate to another site). This is a question for future work.

Our experiments also directly revealed additional important boundary conditions. First, as documented in the pooled analysis, the effect appears to be larger among more expensive products, where customers may have greater hesitancy over whether to make a purchase, and where QS may add yet another hesitation. In this situation, offering a closure-oriented "one-click" choice process can tip customers over the edge to buy.

Second, the pooled analysis revealed a larger quantity integration effect when customers respond to different calls to action in the two formats than when they respond to the same generic question. Leading with a question like, "How many product x, if any, would you like to buy?" may enhance the power of quantity integration, although it is not required in order to observe a benefit from quantity integration.

Third, the field experiment demonstrated that quantity integration is less effective in situations where customers already know they need to buy and where they tend to purchase multiple product complements. In the former case, when customers already know they need to buy, they may have effectively surpassed the purchase decision in QS, rendering the two selling formats equivalent. And in the latter case, the allure of closure—a "one click seals the deal" feeling—may be canceled out by the need to continue considering complements to the desired product. Thus, we suspect that quantity integration should be the most impactful in cases where the decision can feasibly be reached in an efficient manner with a single click, and where customers have not already decided that they wish to buy. However, the field experiment was not designed to tease apart which of these features was the stronger contributor to this boundary condition, so future research is needed to contrast the role of the urgent need to buy and the need to buy multiple complements.

Theoretical Implications and Possible Extensions

A key theoretical contribution of this work is in shedding light on what it means to say "yes." Specifically, we find that customers are more apt to "say yes" and make a purchase when that yes is accompanied by a specific quantity. This suggests that additions to "yes" that provide closure (here, the addition of a specific quantity) can increase consumers' willingness to select it. Moreover, these findings suggest an important distinction between simple yes/no decisions and the cases in which "yes" necessitates

further choices downstream (e.g., choosing quantity, customizing, selecting add-on features, etc.)—that is, where "yes" provides less closure. In doing so, the current research provides a new theoretical lens through which to examine the ubiquitous "yes or no" question.

These results suggest that consumers' willingness to engage in various behaviors can be shaped by slight changes in framing that alter the closure afforded by the option to act. This insight also introduces a mechanism that may contribute to behavior change interventions. Prior research finds that prompting people to develop a plan to perform a behavior increases their likelihood of actually performing it (Gollwitzer 1999; Gollwitzer and Sheeran 2006; Nickerson and Rogers 2010). This literature suggests that such "implementation intentions" work because they trigger an association between a specific moment (e.g., a date and time) and the desired behavior (e.g., going to vote; Gollwitzer and Sheeran 2006). Our findings suggest a possible supporting mechanism: when the behavior appears more precise (i.e., because it includes a date and time), it may offer greater closure, increasing individuals' willingness to agree to perform it.

The mechanism of closure is quite general, and we suspect it may also contribute to other factors that have been shown to encourage choice or purchase. As an example, it is generally believed that one of Amazon's major steps in increasing growth was eliminating the shipping fees for Prime customers (Wei 2018). It is possible that part of that success may be driven by a mechanism analogous to the closure afforded by QI: by eliminating any need to consider shipping costs, Amazon propelled customers closer to a "one click

seals the deal" experience, which may have encouraged them to buy. This may be an interesting topic for future research.

We further propose that the power of quantity integration may apply not only to purchase decisions, but also to other situations in which there is both a decision to act and a choice of quantity. Consider the decisions of whether and how much to invest in the stock market, whether and how much to exercise, whether and how much to volunteer, whether and how much to diet, whether and how much to study for a test, and so on. In all of these cases, simultaneously considering the behavior and its quantity may increase the closure afforded by the option to act, and accordingly may increase individuals' likelihood of choosing to do so. We hope our findings open the door to exploring not only whether, but how many, such contexts reveal the power of quantity integration.

ACKNOWLEDGEMENTS

We are grateful to Jonathan Levav and Andrea Morales for helpful comments on this paper, and to Wendy Liu, Ayelet Gneezy, and Piotr Winkielman for feedback on the ideas discussed herein.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Duke, Kristen Elizabeth and On Amir. The dissertation author was the primary investigator and author of this paper.

FIGURES

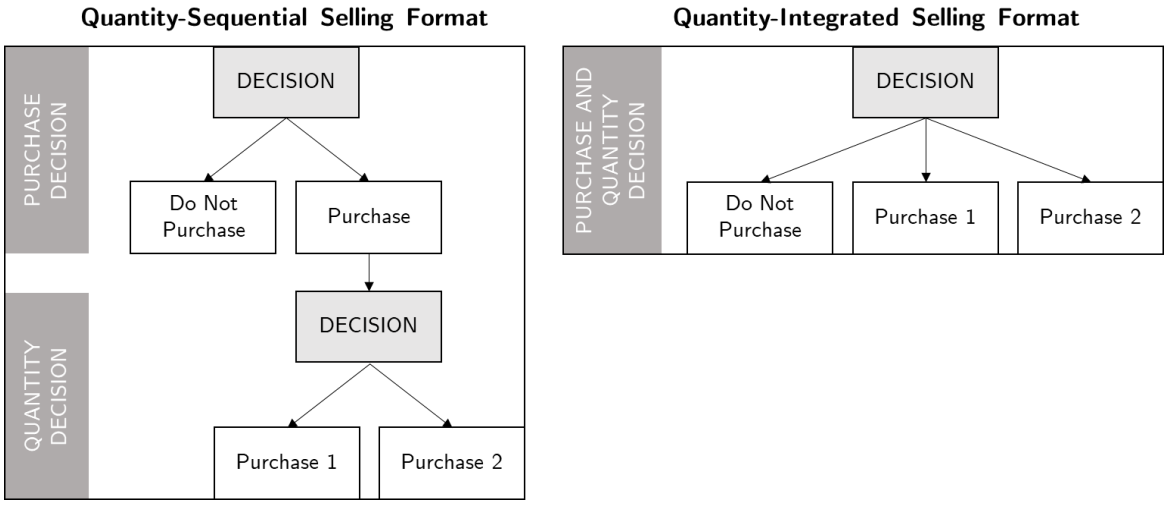


Figure 2.1. Schematic of QS and QI selling formats. This image depicts a typical QS format where customers first choose to purchase and then choose the quantity. Note that a QS format could also require customers to first consider quantity and then evaluate whether to purchase; what defines QS is simply that these are two distinct decisions. In contrast, what defines the QI format is the requirement to simultaneously respond to both questions in a single choice.

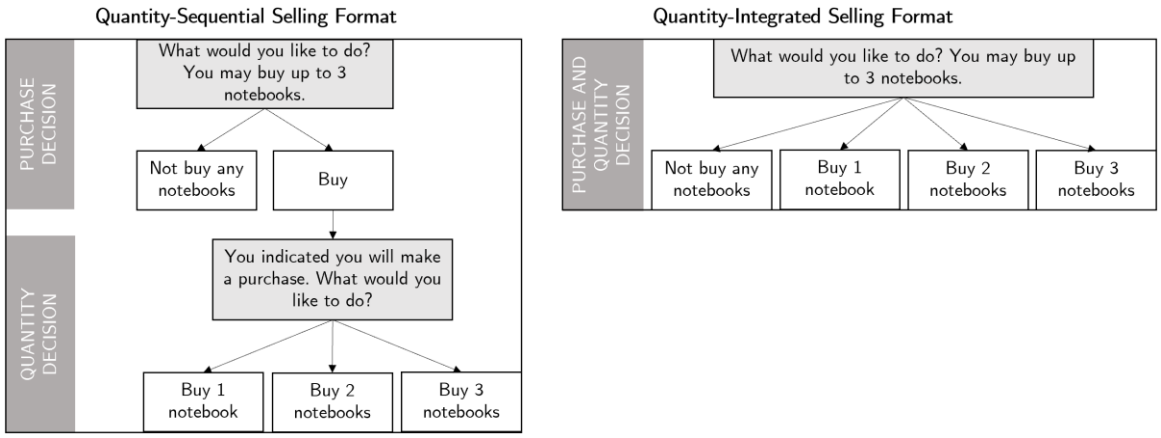


Figure 2.2. Theoretical implementation of QS and QI selling formats.: In a theoretical implementation, customers face the same call to action (respond to the same question), receive the same information (including the maximum purchase quantity), and see the same non-purchase option in both selling formats.

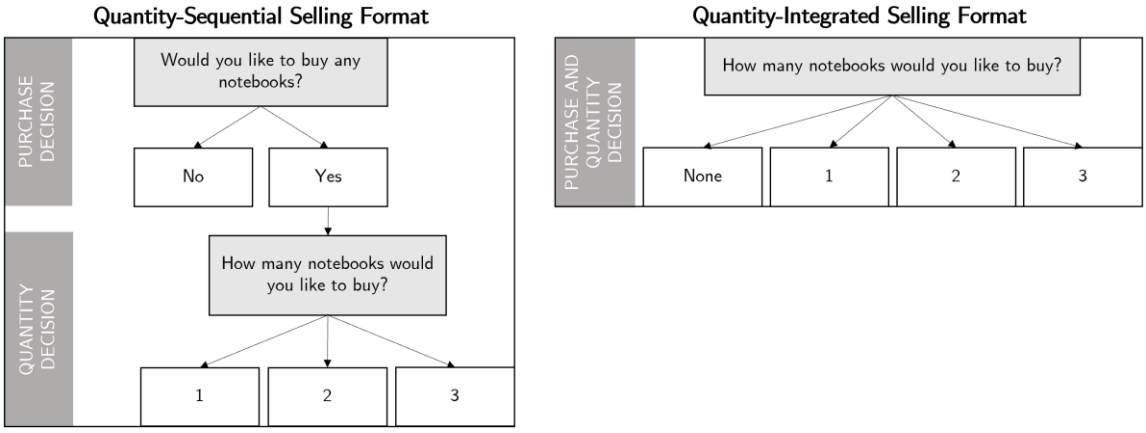


Figure 2.3. Practical implementation of QS and QI selling formats.

QS Screen **QI Screen**

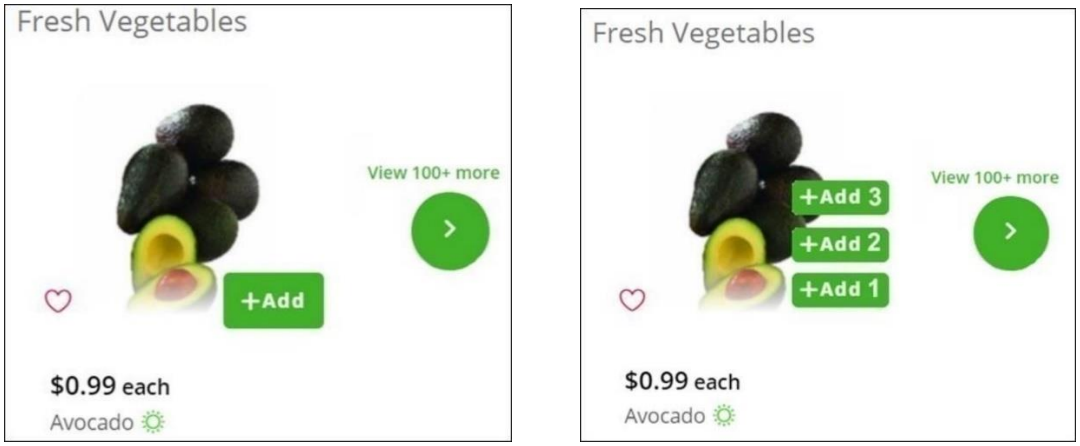


Figure 2.4. Choice screen for each selling format, Experiment 1B.

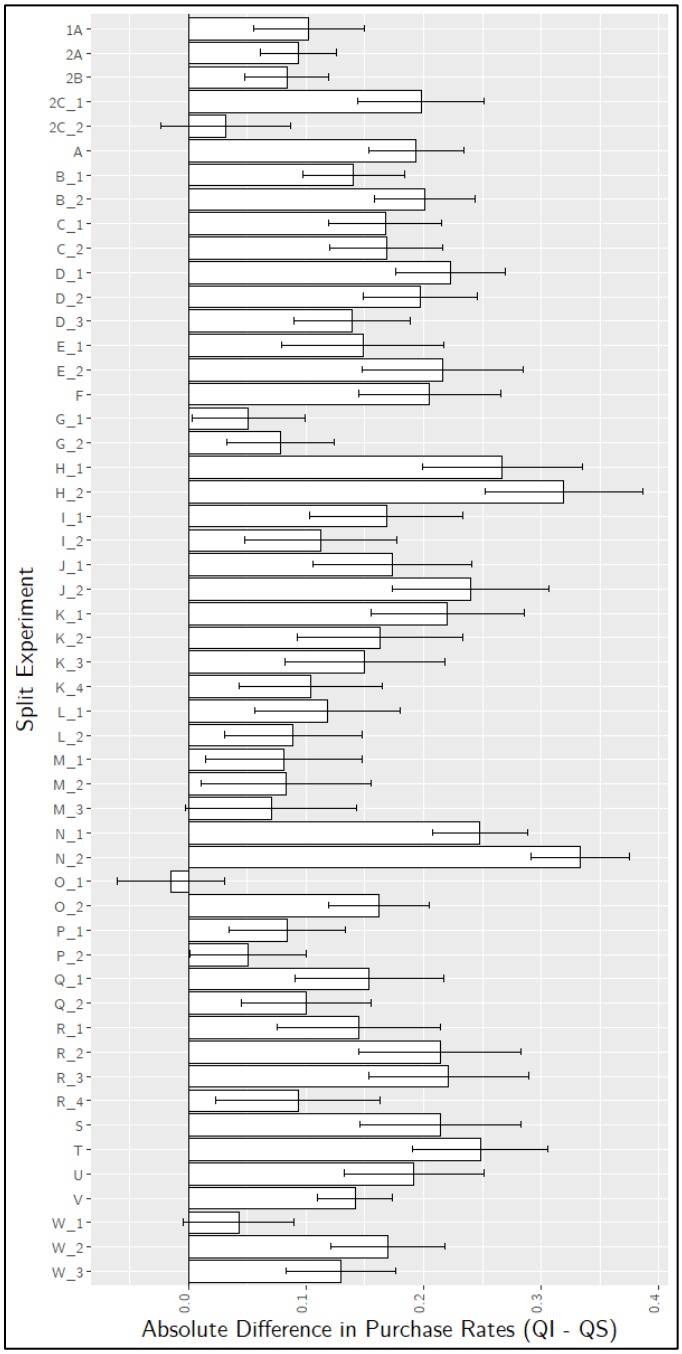


Figure 2.5. Difference in purchase rate between selling formats across split experiments. Error bars represent ±1 SE of the difference between formats.

Figure 2.6. Schematic of QS finder interface.

Figure 2.7. Schematic of QI finder interface.

TABLES

Table 2.1. Logistic regressions and linear probability models predicting purchase, all experiments.

	generalized linear mixed-effects			*linear mixed-effects*		
	(1)	(2)	(3)	(4)	(5)	(6)
Selling Format (SF)	.80***	.80***	.84***	.15***	.15***	.15***
	(.04)	(.04)	(.17)	(.01)	(.01)	(.03)
Mentioned Max Q		.81**	.84**		.15**	.16**
		(.26)	(.27)		(.05)	(.05)
Max Quantity		.07	.06		.01	.01
		(.07)	(.07)		(.01)	(.01)
Same Question		−.31	−.15		−.06	−.03
		(.24)	(.25)		(.05)	(.05)
Existing Purch		−.31	−.28		−.06	−.05
		(.23)	(.24)		(.05)	(.05)
Price		−.21*	−.24**		−.04*	−.05**
		(.08)	(.09)		(.02)	(.02)
SF x Mentioned Max Q			−.03			−.01
			(.10)			(.02)
SF x Max Q			.02			.01
			(.03)			(.01)
SF x Same Question			−.31**			−.05*
			(.10)			(.02)
SF x Existing Purch			−.05			−.02
			(.08)			(.02)
SF x Price			.07*			.01*
			(.03)			(.01)
Constant	−.57***	−1.05*	−1.08*	.39***	.30***	.29**
	(.12)	(.45)	(.45)	(.02)	(.09)	(.09)
Observations	16,431	16,431	16,431	16,431	16,431	16,431
Akaike Inf. Crit.	20,461.94	20,459.36	20,431.45	21,376.56	21,400.48	21,412.59
Bayesian Inf. Crit.	20,492.77	20,528.73	20,539.34	21,415.10	21,477.55	21,528.19
Note:						*p<0.05; **p<0.01; ***p<0.001

APPENDIX

EXAMPLES OF RETAILER USE OF THE TWO SELLING FORMATS

Examples of The Quantity-Sequential Selling Format

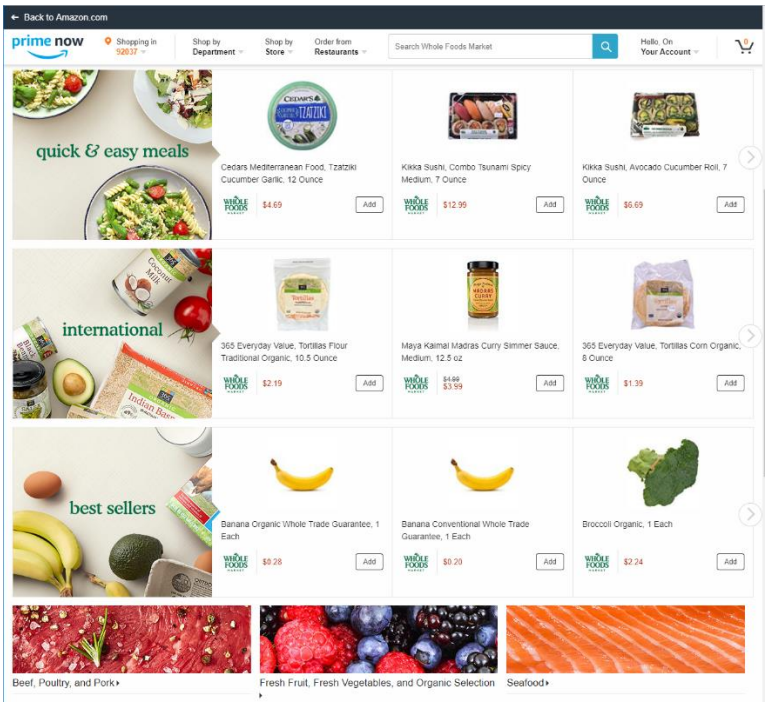


Figure A2.1. Amazon's display of grocery items available on "PrimeNow". Note that customers are only given opportunities to "add" an item to the cart, without simultaneously specifying quantity.

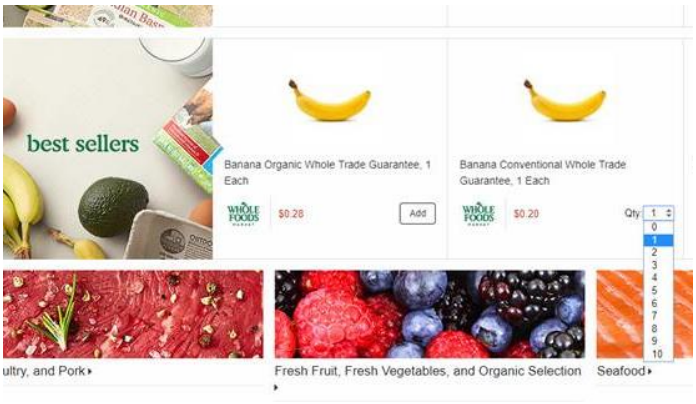


Figure A2.2. After a user clicks the "Add" button next to a product offer, he/she is given an opportunity to specify/modify the quantity using a drop-down menu.

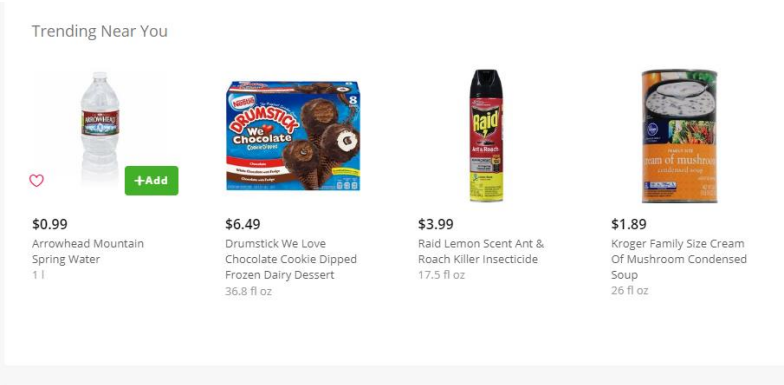


Figure A2.3. Instacart's interface options upon hovering over a focal product. As with PrimeNow, customers here are only given opportunities to "+Add" an item to the cart, without simultaneously specifying the purchase quantity.

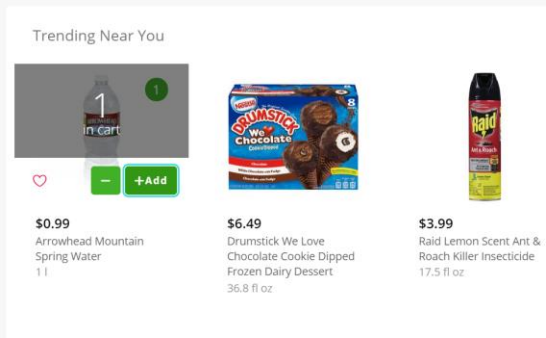


Figure A2.4. After a user clicks the "+Add" button, he/she is given an opportunity to specify/modify the particular quantity by clicking the button again or clicking the minus button.

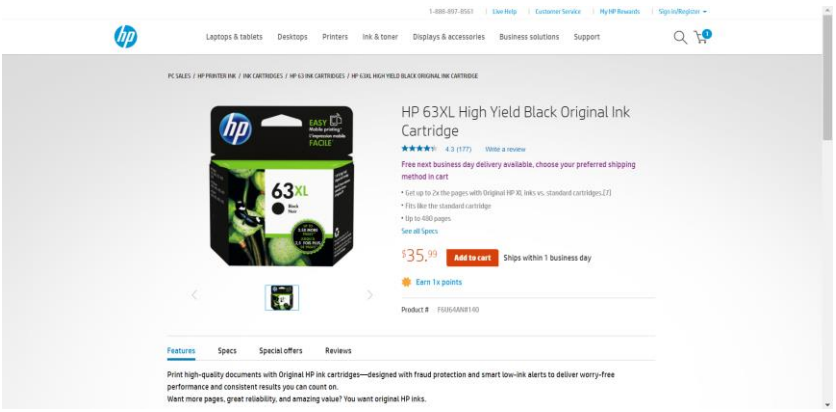


Figure A2.5. HP's product specifics page. Customers here can only choose to "Add to cart" and thereafter can adjust the quantity via a drop-down menu in the shopping cart.

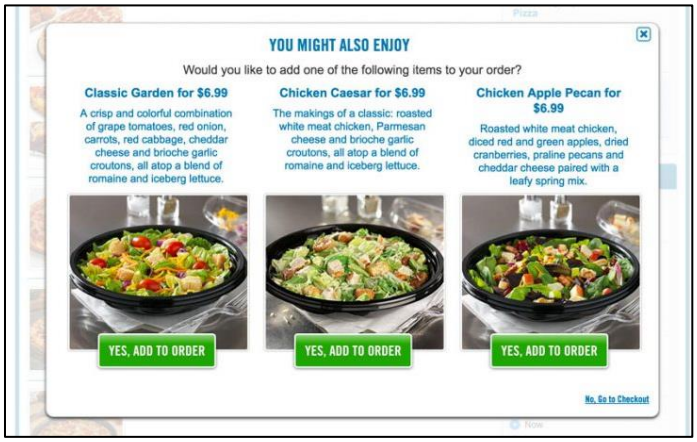


Figure A2.6. Domino's item recommendations in a pop-up message. Clicking adds one unit of the item to the user's shopping cart; the user can then adjust the quantity on the shopping cart page:

Examples of The Quantity-Integrated Selling Format

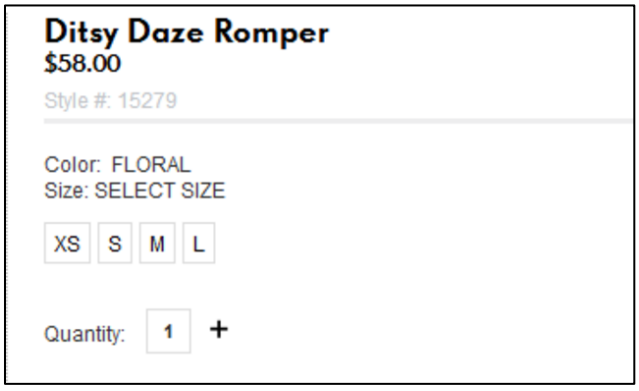


Figure A2.7. Quantity selected by clicking an arrow.

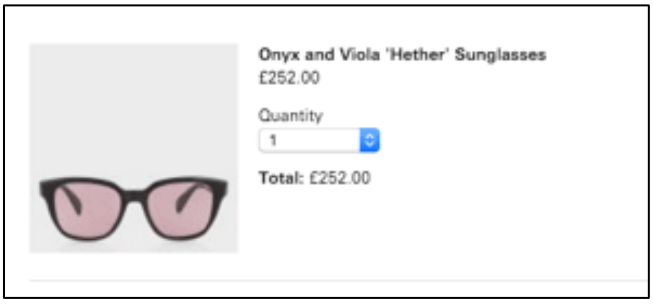


Figure A2.8. Quantity selected via a drop-down menu.

EXPERIMENT MATERIALS

Experiment 1A

Participants saw Figure A2.9 when making their purchase decisions.

If you are selected, you will receive a **$20 bonus**.

If you'd like, you can use this money to purchase Lemome notebooks (pictured below) for **$6 per notebook** (current price on Amazon: $9). **You can buy up to 3 notebooks.**

Any money you do not spend on notebooks will be given to you as a bonus.

Figure A2.9. Experiment 1A Materials.

Experiment 1B

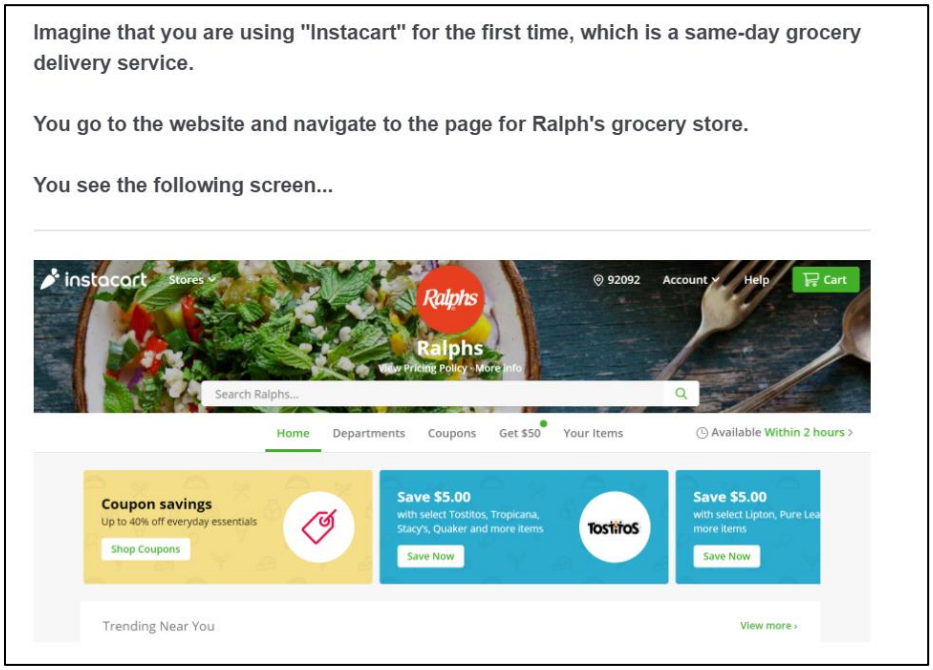


Figure A2.10. First, participants saw the landing page on the website.

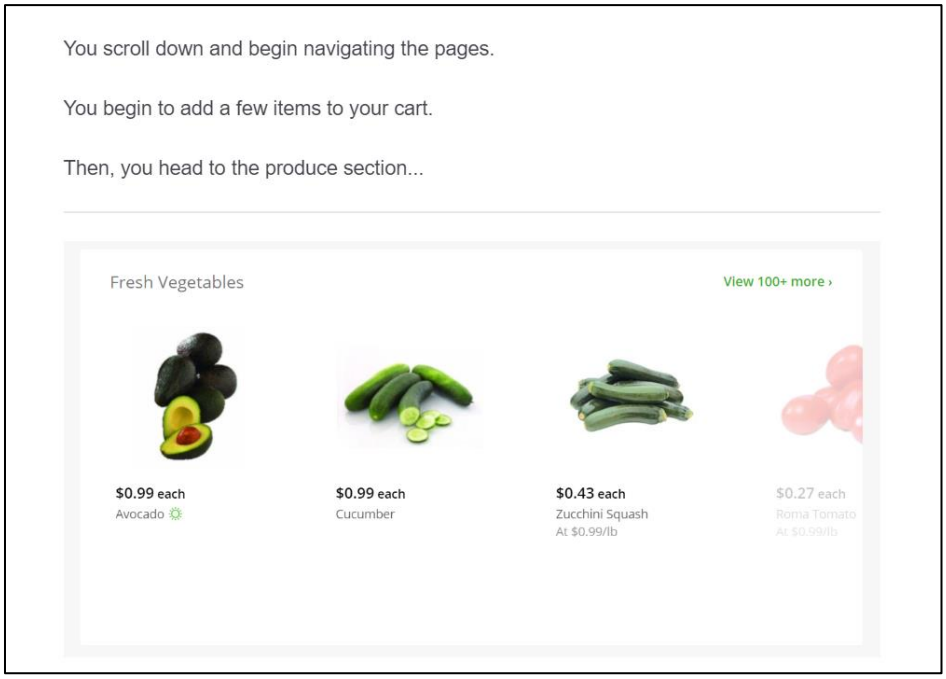


Figure A2.11. Then, they were led to the Produce section of the site.

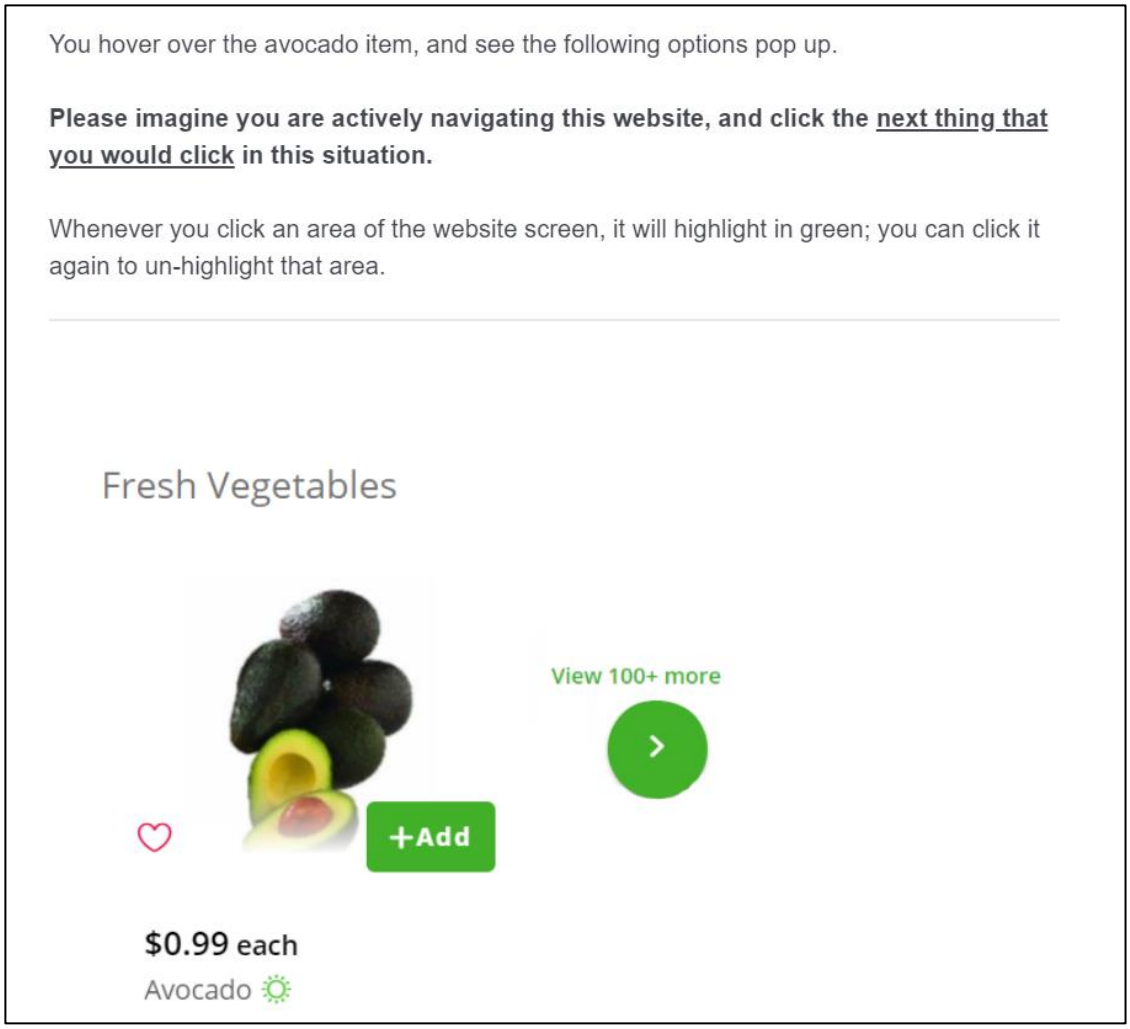


Figure A2.12. Finally, they saw the focal page in which they made a "decision" (i.e., chose what to click next). Note that the available choice options varied by selling format, as shown in the manuscript (the screenshot below represents the QS condition).

Experiment 2A

Participants saw Figure A2.13 when making their purchase decisions.

If you are selected, you will receive a **$10 bonus**.

If you'd like, you can use this money to purchase bags (6 oz. each) of Lindt milk chocolate truffles for **$2.50 per bag.** You can buy up to 4 bags.

Any money you do not spend on truffles will be given to you as a bonus.

Figure A2.13. Experiment 2A Materials.

Experiment 2B

Participants saw Figure A2.13 before making a choice.

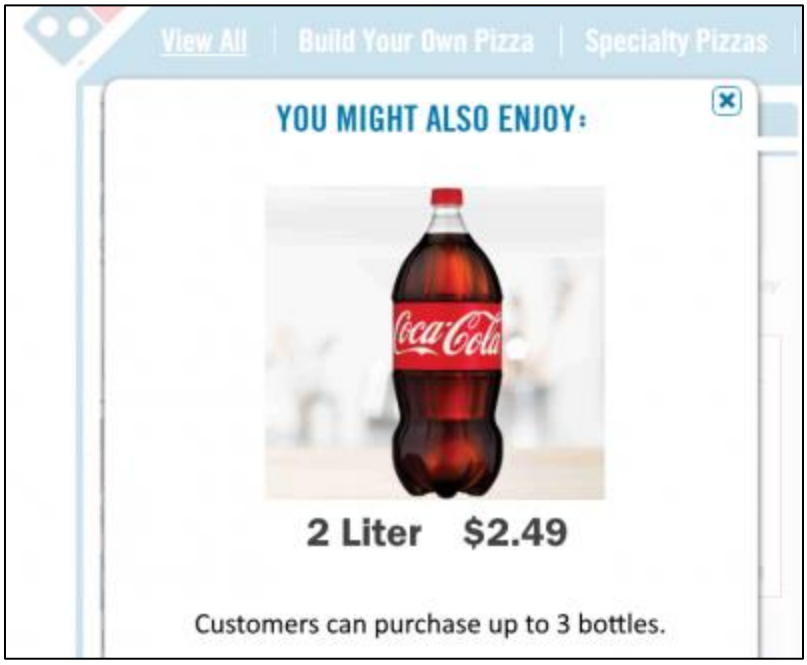


Figure A2.14. Experiment 2B Materials.

Experiment 2C

Participants saw Figure A2.14 when making their purchase decisions:

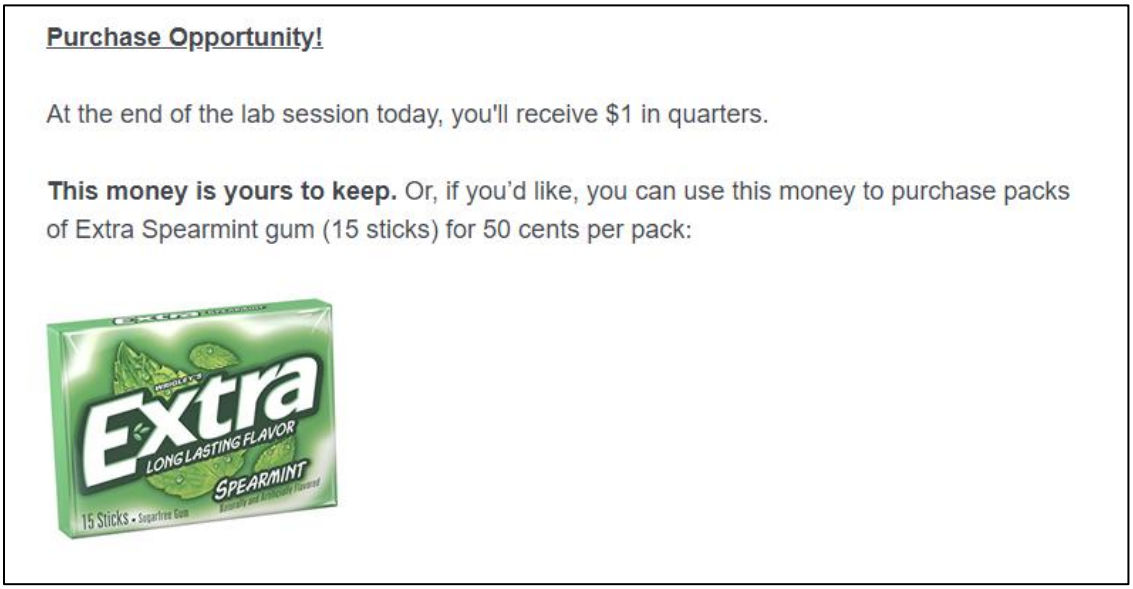


Figure A2.15. Experiment 2C Materials.

ADDITIONAL EXPERIMENTAL DETAILS FOR EXPERIMENTS 2A—2C

Experiment 2A

Additional Method Details

As mentioned in the paper, after making their initial purchase choice, all participants responded to three measures capturing their sense of closure with their choices. These measures were adapted from prior research investigating choice closure (Gu, Botti, and Faro 2013). In randomized order, participants answered, "To what extent do you feel a sense of closure with your decision?," "To what extent are you still preoccupied with your decision?" (reverse-coded), and "To what extent do you perceive your decision as a 'closed book'?" (all 0: Not at all to 100: Completely).

Additional Results Details

We regressed the composite closure measure on the selling format and whether or not participants had purchased. As expected, purchasers reported experiencing less closure than non-purchasers ($\beta = -19.78$, $t(797) = -12.93$, $p < .001$). Furthermore, consistent with our theorizing, QS participants reported less closure than QI participants ($\beta = -4.60$, $t(797) = -3.24$, $p = .001$).

Importantly, the difference in closure across formats was particularly pronounced among participants who had made a purchase (QI: M = 70.78, SD = 22.16 vs. QS: M =

62.71, SD = 22.18; t(248) = 2.84, p = .005). Non-purchasers also reported marginally greater closure in the QI format (M = 88.26, SD = 18.61) than in the QS format (M = 85.21, SD = 19.04; t(548) = 1.90, p = .06). To analyze further, we regressed closure on format (QS vs. QI, effect-coded [i.e., 1 vs. -1]), purchase (purchased vs. did not purchase, effect-coded), and their interaction. Consistent with the previous results, this analysis revealed a significant main effect of format (β = -2.78, t(796) = -3.62, p < .001), whereby QS participants reported lower closure, and a significant main effect of purchase (β = -10.00, t(796) = -13.03, p < .001), whereby non-purchasers report higher closure than purchasers. Further, these effects were qualified by a marginal interaction (β = -1.25, t(796) = -1.63, p = .10), whereby the difference in closure across formats was larger among purchasers than among non-purchasers (see Figure A2.15).

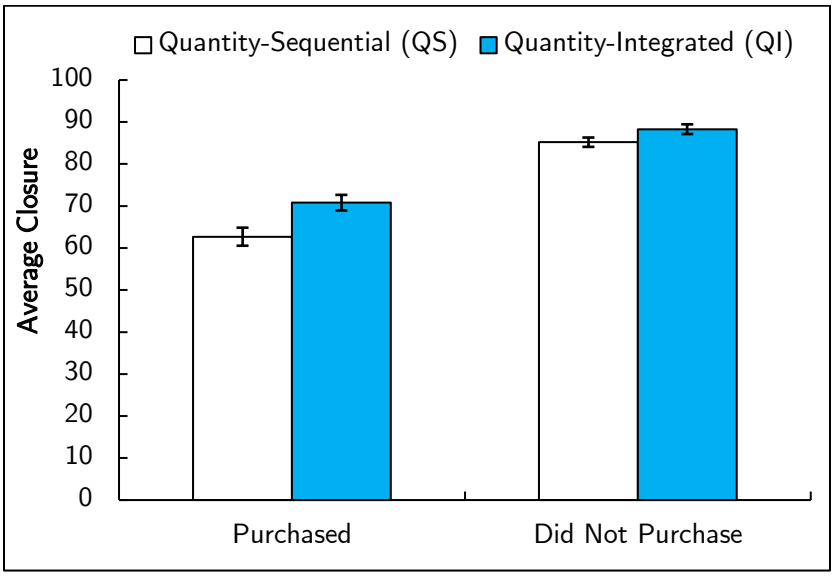


Figure A2.16. Average sense of closure in each selling format among purchasers and non-purchasers, Experiment 2A. Error bars represent ±1 SE around the mean.

Experiment 2B

Additional Method Details

In this experiment, participants also answered an attention check question: "What was the product you just made a decision about?" (with options: pizza, hamburgers, vitamin water, Coca Cola, Sprite, chocolate chip cookies, M & Ms, and I'm not sure/I don't remember). For consistency with other studies that do not include an attention check, we retained all participants regardless of their response; excluding participants who failed this check yields the same results.

As mentioned, participants completed the 15-item abridged Need for Closure Scale (Roets and Van Hiel 2011). This scale includes items such as, "I don't like situations that are uncertain," "When I am confronted with a problem, I'm dying to reach a solution very quickly," and "When I have made a decision, I feel relieved" (1: Completely disagree to 6: Completely agree).

Additional Results Details

In the main paper, we include the covariate of Coca-Cola purchasing frequency because when examining an individual difference variable, not doing so could introduce a potential confound. For completeness, the results without the covariate are: a significant main effect of selling format (β = -.22, SE = .09, z = -2.35, *p* = .02), and no main effect of NFCC (β = -.10, SE = .13, z = -.73, p = .46), qualified by a marginal NFCC x selling format interaction (β = -.24, SE = .13, z = -1.91, p = .06). Thus, they converge with the reported results. See figure A2.16 for a graphic representation of this pattern.

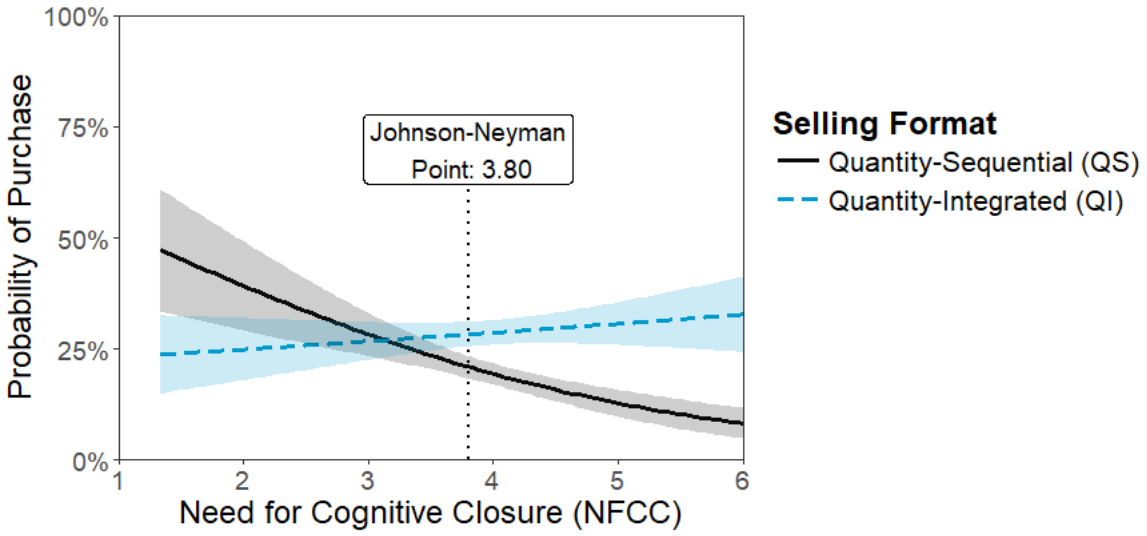


Figure A2.17. Probability of purchasing in each format, as a function of NFCC, Experiment 2B. The shading represents ±1 SE around the predictions.

Quantity purchased. Among those who purchased, most participants (81.76%) purchased exactly one bottle. Purchasers bought similar numbers of bottles in both formats (QS: M = 1.27, SD = .59; QI: M = 1.22, SD = .51; t(157) = -.58, p = .56). To test whether NFCC affected the quantity purchased, we regressed purchase quantity (conditional on purchase) on the selling format (effect-coded), NFCC (mean-centered), and their interaction, with Coca Cola purchasing frequency as a covariate. This analysis revealed no main effects nor interactions (p's > .55; the same result arises when excluding the covariate). Thus, NFCC did not impact purchase quantity.

Total sales. The QS condition sold fewer total bottles of soda (85; N = 302) than did the QI condition (112; N = 301): changing from a QS to a QI format increased total sales by 32%.

Experiment 2C

Pretest

We conducted a pretest to verify the priming manipulation in this experiment. Three hundred Mechanical Turk workers (48.3% female; M_{age} = 37.0 years, SD_{age} = 11.8 years) participated. Participants were randomly assigned to one of three conditions (prime: lack of closure, high sense of closure, or control) in a between-subjects design.

Prime. Participants in the lack of closure condition read and completed the same task that Experiment 3 participants in the "lack of closure" condition completed: "For the next two minutes, please think of something that is ongoing in your life, where the outcome is yet to be determined. Focus on how this thing is not yet complete and still an unpredictable, "open book." In the space below, please describe how this makes you feel." Those in the high sense of closure condition read and completed the same task that Experiment 3 participants in the "high sense of closure" condition completed: "For the next two minutes, please think of a book you recently read or a movie you recently watched. In the space below, please describe the basic premise of this book or movie." Those in the control condition skipped this prime phase and proceeded to the key measures assessing feelings of lacking closure.

Feelings of Lacking Closure. All participants next read, "Please think about how you are feeling right now. Then, answer the following about your current feelings." They then responded to six items (modified from Gu, Botti, and Faro 2013 and Webster and Kruglanski 1994) designed to capture their feelings of lacking closure—that is, the feeling that we propose underlies the quantity integration effect. In randomized order, participants answered, "Right now I feel unsettled," "Right now, I feel a sense that things are resolved" (reverse-coded), "Right now, I have a satisfying feeling of completeness" (reverse-coded), "In this moment, I feel a sense of closure" (reverse-coded), "Right now, I feel like something is nagging away at me," and "Based on how I'm currently feeling, I do not want to be in an unpredictable situation right now" (all 1: Strongly Disagree to 7: Strongly Agree). Thereafter, participants in the control condition completed the high closure priming manipulation to equate time spent on the survey across conditions. Finally, all participants provided demographic information.

Pretest Results

Factor analysis. To identify the underlying structure of the self-report measures we collected (as they were modified from the original source items), we conducted a principal components factor analysis with Varimax rotation. This analysis revealed that the items loaded onto two distinct factors, together explaining 70% of the variance. Factor 1 (the Feeling of Lacking Closure; four items) explained 51% of the variance and captured our

primary concept of interest. Factor 2 (Nagging or Unpredictable; two items) explained 20% of the variance and accounted for the two other measures (having something nagging away at oneself and not wanting to be in an unpredictable situation). Reliability analyses indicated high reliability for the Feeling of Lacking Closure factor (α = .89) but not for the Nagging and Unpredictable factor (α = .29). Thus, we focused our attention primarily on the first factor. See Table A2.1 for all items and factor loadings.

Table A2.1. Factor analysis of self-report measures.

	Feeling of Lacking Closure	**Nagging or Unpredictable**
Feeling of Completeness (R)	.92	
Things are Resolved (R)	.91	
Sense of Closure (R)	.89	
Feel Unsettled	.75	
Something Nagging Away		.78
Averse to Unpredictable		.76

Feeling of Lacking Closure. We expected that our lack of closure prime would induce a high feeling of lacking closure (relative to control), but that our high sense of closure prime would attenuate this feeling (relative to control). Results supported this prediction. Participants in the lack of closure priming condition reported significantly higher feelings of lacking closure (M = .71, SD = .92) than did those in the control condition (M = -.21, SD = .92, $t(195) = 7.03$, $p < .001$) and those in the high sense of closure condition (M = -.51, SD = .72, $t(202) = 10.54$, $p < .001$). Furthermore, those in the control condition reported significantly higher feelings of lacking closure than did those in the high sense of closure condition ($t(197) = 2.58$, $p = .01$). Thus, relative to participants' natural baselines, the lack of closure prime amplified the feeling of lacking closure, while the high sense of closure prime attenuated this feeling.

Nagging or Unpredictable. Participants' responses did not significantly differ on this factor. Those in the lack of closure condition were marginally higher on this factor (M = .08, SD = .91) than were those in the control condition (M = -.13, SD = .76, $t(195) = 1.79$, $p = .07$). Those in the high sense of closure condition (M = .04, SD = 1.25) did not differ from control participants ($t(197) = 1.20$, $p = .23$) nor the lack of closure participants ($t(202) = .24$, $p = .81$). Thus, the priming manipulation primarily altered participants' responses to the first factor.

Experiment: Additional Method Details

The priming manipulation was as follows. All participants began by completing a two-minute writing task designed to prime either a sense of lacking closure or a high sense of closure (see full text. Participants in the lacking closure condition read, “For the next two minutes, please think of something that is ongoing in your life, where the outcome is yet to be determined. Focus on how this thing is not yet complete and still an unpredictable "open book." In the space below, please describe how this makes you feel.” This manipulation was designed to induce the sense that things are incomplete or unfinished—that is, the sense of lacking closure.

Participants in the high sense of closure condition instead read, “For the next two minutes, please think of a book you recently read or a movie you recently watched. In the space below, please describe the basic premise of this book or movie.” This manipulation was designed to induce the feeling that things are complete and resolved—that is, the sense of having closure—based on the notion that most literary and film pieces are intentionally designed to offer a satisfying sense of completeness and resolution by the end (Miller 1978; Adamo 2000; Richter 1978). Note that after the priming manipulation, participants also completed a set of measures tapping their beliefs about general (un)certainty in life. Specifically, they answered, “In general, to what extent do you perceive life as: uncertain, stable, predictable, surprising, incomplete, unknown” (randomized; all items on a scale from 1: Strongly disagree to 7: Strongly agree). Those in the lacking closure condition reported perceiving life as significantly more incomplete than those in the high sense of closure condition ($p = .01$) but did not significantly differ on the other measures. Controlling for these perceptions does not influence the main results.

Additional Results Details

Figure A2.18 graphically presents the moderation described in the main paper.

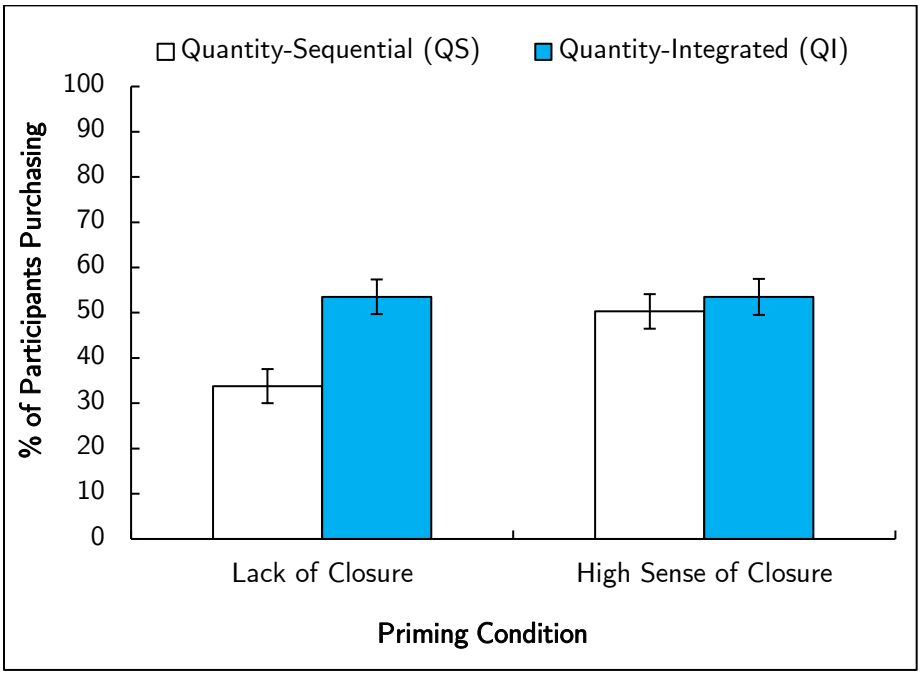


Figure A2.18. Purchase as a function of format and the priming manipulation, Experiment 2C.

Quantity purchased. Purchasers bought similar quantities in both formats (QS: M = 1.71, SD = .46; QI: M = 1.65, SD = .48; t(310) = -1.06, p = .29). To assess the effect of the prime on quantity purchased, we regressed quantity purchased (conditional on purchase) on selling format (QS vs. QI, effect-coded), prime (lack of closure vs. high sense of closure, effect-coded) and their interaction. This analysis revealed no main effects and no interaction (p's > .26).

Total sales. The QS condition sold fewer total packs of gum (234; N = 328) than did the QI condition (289; N = 327): changing from a QS to a QI format increased total sales by 25%.

\DETAILS ON ALL EXPERIMENTS

Methods for All Experiments

All of our lab experiments follow a similar structure[10]. We list the details of all experiments in Supplemental Table 1 on the following page, including those in the main paper (represented by numbers, e.g., 1) and those not in the main paper (represented by letters, e.g., A). Column 2 lists the number of participants; unless mentioned otherwise, all

[10] Note that Experiment 1B in the paper is the only exception; it follows a different paradigm and therefore is not included in this analysis.

participants were Mechanical Turk workers. Participants were assigned to condition in a between-subjects design; Column 3 outlines the factorial design of each experiment.

All participants read a scenario in which they imagined being in a shopping situation or actually entered a consequential purchase situation. Column 4 presents the scenario or purchase opportunity description. Then, to provide greater context, all participants viewed an image of a pop-up advertisement, a sign advertising a sale, or an image of the product available for purchase. The product and its price (and whether the price was described as a discounted price) are in Columns 5 and 6. Column 7 presents the quantity limit (i.e., the maximum number of units participants they could purchase), and notes whether or not this information was provided to them explicitly in both selling formats. All participants then made a purchase decision: they responded to a question by selecting a choice option, according to their selling format. Columns 8 and 9 present these questions and choice options.

Several experiments were designed with a secondary goal beyond simply demonstrating the effect in a new setting. Accordingly, we often orthogonally layered an additional manipulation on the design. In some cases, the results of this additional manipulation yielded meaningful insights about potential psychological mechanisms; for these experiments, we provide a concise description of the methods and results later. In other cases, the results were not relevant for interpreting theory; for these experiments, we simply footnote summary results information in Supplemental Table 1.

Results for All Experiments

As mentioned previously, Supplemental Table 1 includes several experiments that included an additional, orthogonal manipulation beyond the manipulation of selling format. Accordingly, as described in the main manuscript, we split each such experiment into separate smaller experiments, each with only two conditions: the QS selling format or the QI selling format.

Supplemental Table 2 presents the percentage of participants purchasing in each selling format, within each such split experiment. Each experiment that was split is represented with a value in parentheses (e.g., Experiment 2 (I)) and is described in Column 2. Columns 3 and 4 present the percentage of participants purchasing in each selling format, and Column 5 presents the χ^2 statistic analyzing the difference between the two selling formats and notes the level of significance.

SINGLE-PAPER META-ANALYSIS

Following the procedures outlined by McShane and Böckenholt (2017), we performed a single-paper meta-analysis. This revealed a significant effect of selling format

on participants' likelihood of purchasing across all (split) experiments. According to this analysis, the QI selling format led to a 15.05 percentage-point increase (.1505 effect) in likelihood of purchase relative to the QS format (95% CI [12.97%, 17.13%], *SE* = .01 (1.06%), $z = 14.20$, $p < .001$).

The point estimates were a 38.74% average purchase rate in the QS condition (*SE* = 2.35%) and a 53.79% average purchase rate in the QI condition (*SE* = 2.35%). Thus, this tool estimates the relative increase in purchase rates associated with quantity integration at 39%. This analysis confirms a robust effect of selling format. Figure A2.18 shows the magnitude of the effects of selling format in each split experiment.

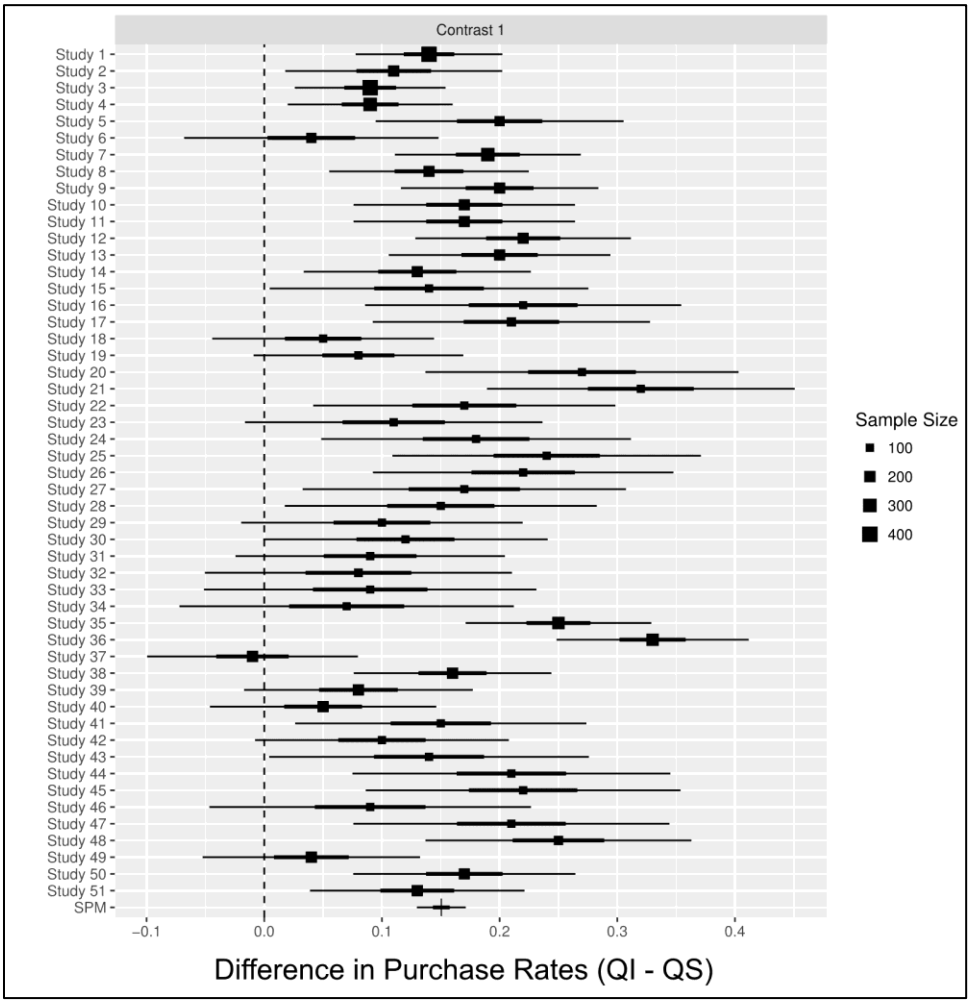


Figure A2.19. Single-Paper Meta-Analysis (https://blakemcshane.shinyapps.io/spmeta/). The point estimated size of the quantity integration effect are presented in squares, and this difference is represented on the x-axis. 50% and 95% confidence intervals of the magnitude of the effect are presented with the thick and thin lines, respectively.

This meta-analytic tool also revealed that I^2 was estimated at 95.84% (95% CI [95.35%, 96.29%]), suggesting a substantial level of heterogeneity. Thus, method factors (e.g., the operationalization of the experimental manipulation, or unaccounted-for moderators in this analysis such as the product) contribute to a high percentage of the variation in observations.

SUMMARY WRITEUP OF FOCAL APPENDIX EXPERIMENTS

Experiment I: Adding a Deferral Option

Participants imagined encountering a pop-up advertising a sale on bar soap. QS participants indicated whether or not they wanted to buy before choosing quantity, while QI participants directly indicated the quantity (with a 0 default). Half of participants saw an additional option to defer: "I want to decide later." QS participants were again significantly less likely to purchase than were QI participants, regardless of the deferral opportunity (24.38% vs. 38.38%; χ^2 (1, N = 399) = 9.09, p = .003). There was no main effect of the availability of a deferral option (β = -.04, z = -.37, p = .71) and no interaction (β = -.06, z = -.56, p = .58). However, QS (vs. QI) participants were also significantly more likely to defer the choice when given the chance to do so (16.83% vs. 3.00%; χ^2(1, N = 201) = 9.24, p = .002), suggesting that non-purchases in QS choices may not necessarily represent a true desire to not purchase, but in some cases, may reflect a decision centered on uncertainty and avoidance of the choice to purchase.

Experiment R: Varying the Maximum Quantity

Participants imagined ordering a pizza online and encountering a pop-up advertising a sale on Coke. They could purchase a maximum of 1, 3, or 5 Coke bottles (between-subjects). QS answered, "Would you like to add any 2 liter bottles of Coke to your order?" with either "yes" or "no." QI participants answered, "How many 2 Liter bottles of Coke would you like to add to your order?" with options "0" and "1" (when the maximum quantity was 1, and "2" and "3" as additional options when the maximum quantity was 3; etc.). Significantly more QI (57%) than two-stage (40%) participants purchased soda (χ^2(1) = 22.94, p < .001). Importantly, this effect was significant at all maximum quantity values (p's < .05), including one, where there were two choice options in both formats.

Experiment V: QI Layout as Sliding-Scale Versus Drop-Down Menu

Participants imagined ordering a pizza online and encountering a pop-up advertising a sale on Coke. QS participants answered, "Would you like to add any 2 liter bottles of Coke to your order?" with either "yes" or "no." QI participants instead answered, "How many 2 Liter bottles of Coke would you like to add to your order?" on either a

sliding scale or a drop-down menu, with options: 0 (the default), 1, 2, 3, 4, and 5. QS participants were significantly less likely to purchase (22%) than both QI slider participants (38%; $\chi^2(1) = 17.50$, $p < .001$) and QI drop-down participants (34%; $\chi^2(1, = 9.04$, $p = .003$); the two QI formats did not differ ($\chi^2(1) = .86$, $p = .35$).

Experiment W: Experiment Testing Mindset and Construal Level Mechanisms

This experiment was designed to be a well-powered test of two potential explanations for the quantity integration effect. First, when presented with a QI selling format, consumers might initially consider which quantity they prefer, and only thereafter decide whether or not to purchase (e.g., Xu and Wyer 2007). This initial quantity consideration could induce an implemental mindset or stronger implementation intentions (e.g., Gollwitzer and Bayer 1999; Gollwitzer 1999), which would increase purchase likelihood (Dhar, Huber, and Khan 2007). To address this possibility, this experiment tests whether inducing an implemental mindset attenuates the quantity integration effect. Second, it is possible that the formats differ in the level of construal (Trope and Liberman 2003) they invoke, which may contribute to the effect. We test this possibility by assessing participants' momentary construal levels after their choices.

Method

One thousand, two hundred ten Mechanical Turk workers (50% female; M_{age} = 36.8 years, SD_{age} = 12.3 years) participated. Participants were randomly assigned to one of six conditions in a 2 (selling format: QS vs. QI) X 3 (prime: implemental, deliberative, control) between-subjects design.

All participants began by completing a task designed to prime one of three mindsets: implemental, deliberative, or control (no specific mindset). We adapted this priming procedure from Dhar, Huber, and Khan (2007)'s Study 4. As in that study, participants were asked to imagine that they were thinking about buying a car. Then, those in the deliberative condition then wrote down three pros and three cons of buying a car, and those in the implemental condition instead wrote down six steps they would need to take to buy a car. Those in the control condition instead wrote down six things they could do with a car. We expected that most participants would not actively be currently concerned with how to buy a car, and therefore expected that the control condition would match the deliberative condition (consistent with prior research, e.g., Henderson, De Liver, and Gollwitzer 2008).

Thereafter, participants responded to a manipulation check (adapted from Brandstätter and Frank 2002): "Imagine you are about to buy a new car. Do you have a clear sense of what needs to be done to make this happen?" (1: Not at all to 9: Completely).

On the following page, all participants read a scenario description, "Imagine you are online shopping at home. After some browsing, the following box pops up." An image displayed a pop-up with the words, "Scented Candle Sale! ~~$24.50~~ only $6.99!" Text beneath displayed, "in this store, you can buy up to 3 candles per purchase." Then, all

participants answered, “What would you like you do?” For QS participants, the options were “Not buy any” and “Buy.” For QI participants, the options were, “Not buy any,” “Buy 1 candle,” “Buy 2 candles,” and “Buy 3 candles.”

Thereafter, participants completed a 10-item Behavior Identification Form (BIF) questionnaire (Vallacher and Wegner 1989), which has been used to measure participants’ momentary construal level (e.g., Burgoon, Henderson, and Markman 2013; Han, Duhachek, and Agrawal 2016). For each item, participants read a behavior (e.g., “Tooth-brushing”) and choose one of two labels that they feel best describes the behavior—one that is abstract/high-level (e.g., “Preventing tooth decay”) and one that is concrete/low-level (e.g., “Moving a brush around in one’s mouth”). Participants received a score of 0 for each concrete construal they selected, and 1 for each abstract construal description; their responses were summed to form a BIF score.

At this point, QS participants who had previously indicated they would make a purchase selected the quantity. They answered, “How many candles would you buy?” with options 1, 2, and 3. Note that we included this question after the BIF to ensure that the BIF could accurately capture differences in construal level after participants’ initial purchase decisions, which is the point at which we suggest closure diverges. Finally, all participants provided demographic information.

Results

Manipulation check. As expected, participants responded to the priming manipulation check assessing their implemental orientation similarly in the deliberative (M = 7.58, SD = 1.58) and control conditions (M = 7.56, SD = 1.63; $t(816) = -.15$, $p = .88$). Thus, to maximize power in testing for a possible interaction with selling format, we combined these conditions and contrasted them against the implemental prime. Indeed, implemental participants scored significantly higher on the manipulation check (M = 7.78, SD = 1.43) than did deliberative and control participants (7.57, SD = 1.60, $t(1208) = -2.15$, $p = .03$).

Purchase rates. In line with our previous experiments, QS participants were again significantly less likely to purchase any candles (29.8%) than were QI participants (41.0%; $\chi^2(1, N = 1210) = 16.72$, $p < .001$, $\varphi = .12$). To test whether the priming manipulation moderated this quantity integration effect, we performed a binary logistic regression predicting purchase as a function of selling format (QS vs. QI, effect-coded), prime (implemental vs. deliberative/control, effect-coded) and their interaction. This analysis revealed a main effect of selling format ($\beta = .28$, $SE = .07$, $z = 4.31$, $p < .001$), wherein QI (vs. QS) participants were more likely to buy, but no main effect of the prime ($\beta = .03$, $SE = .07$, $z = .45$, $p = .66$), and no significant interaction ($\beta = .09$, $SE = .07$, $z = 1.39$, $p = .17$). This pattern, shown in the Figure A2.19, is directionally opposite from one predicted by the above-mentioned mindset account. These findings demonstrate that implemental mindsets cannot explain the quantity integration effect.

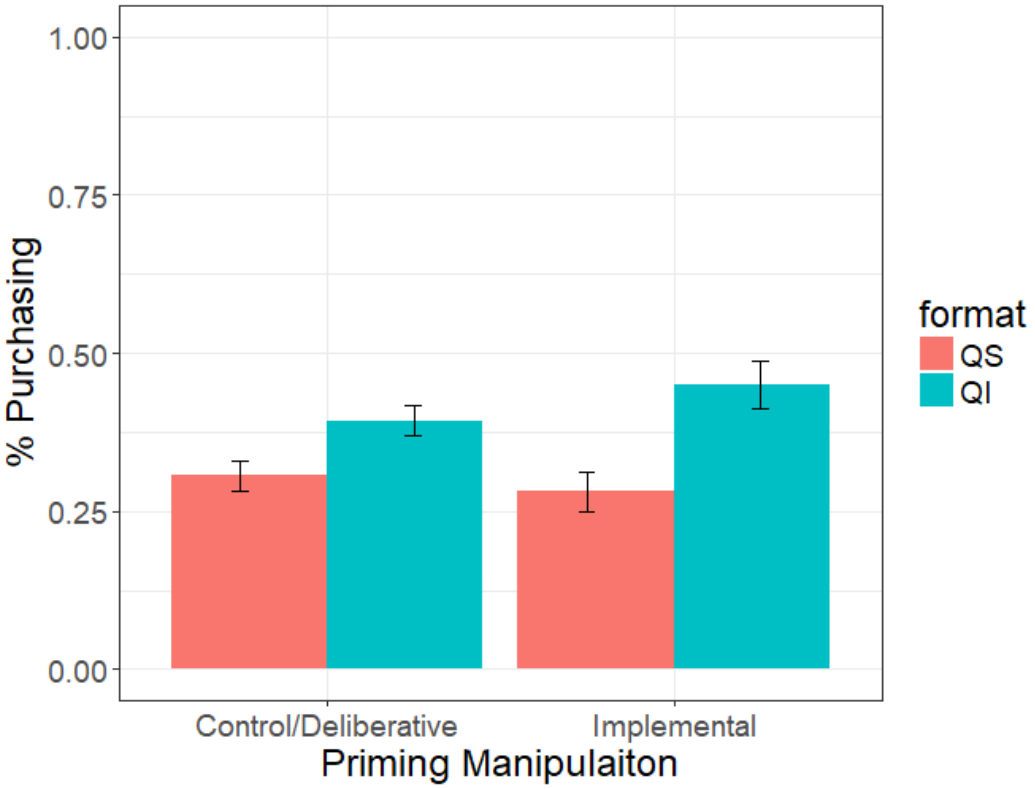


Figure A2.20. Purchase Rates by Format and Priming Condition.

Construal level. We found no differences in construal level between the QS (*M* = 6.42, *SD* = 2.93) and QI formats (*M* = 6.33, *SD* = 2.85; *t*(1208) = -.50, *p* = .62). This persisted regardless of whether or not participants had made a purchase (*p*'s > .54). A regression predicting construal level from selling format (QS vs. QI, effect-coded), purchase (purchased vs. did not purchase, effect-coded) and their interaction revealed no significant effect of format (β = -.03, *SE* = .09, *t*(1206) = -.36, *p* = .72) nor purchasing (β = .002, *SE* = .09, *t*(1206) = .02, *p* = .99), and no interaction (β = .03, *SE* = .09, *t*(1206) = .38, *p* = .71).

Consistent with prior research (e.g., Tu and Soman 2014), and even though a full purchasing scenario separated the two, there was also a link between implemental mindsets and construal level, whereby participants who had initially been primed with the implemental mindset showed a marginally less abstract construal level (*M* = 6.16, *SD* = 3.00) than did control/deliberative participants (*M* = 6.48, *SD* = 2.83; *t*(1208) = 1.83, *p* = .07).

Quantity purchased. In this experiment, among those who purchased, QS participants purchased significantly more candles (*M* = 2.23, *SD* = .82) than did QI participants (*M* = 1.93, *SD* = .88; *t*(426) = -3.57, *p* < .001). A regression predicting the amount purchased among purchasers from selling format (effect-coded), prime (implemental vs. deliberative/control, effect-coded) and their interaction uncovered only

this main effect of format; there was no significant effect of the prime (p = .74) and no interaction (p = .62).

Total sale*s*. Despite selling more units per purchaser, the QS format still sold fewer total candles overall (401; N = 605) than did the QI format (478; N = 605). Changing from a QS to a QI format increased total sales by 19% in this experiment.

Discussion

This experiment demonstrates that implemental mindsets and construal levels do not appear to play a pivotal role in the quantity integration effect. Priming an implemental mindset did not attenuate the quantity integration effect (if anything, it nonsignificantly accentuated it), and the selling formats did not induce different levels of construal.

REFERENCES

Allison, Paul D. (2009), *Fixed Effects Regression Models*, Thousand Oaks, CA: Sage.

Arens, Zachary G. and Rebecca W. Hamilton (2017), "The Rebound of the Foregone Alternative," *Journal of Consumer Psychology*, 27 (3), 318—32.

Beike, Denise R. and Erin T. Wirth-Beaumont (2005), "Psychological Closure as a Memory Phenomenon," *Memory*, 13 (6), 574—93.

Beike, Denise R., Erica Kleinknecht, and Erin T. Wirth-Beaumont (2004), "How Emotional and Nonemotional Memories Define the Self," in *The Self and Memory*, ed. D. R. Beike, J. M. Lampinen, and D. A. Behrend, New York: Psychology Press, 141—59.

Beike, Denise R., Laura P. Adams, and Erin T. Wirth-Beaumont (2007), "Incomplete Inhibition of Emotion in Specific Autobiographical Memories," *Memory*, 15 (4), 375-89.

Beike, Denise R., Keith D. Markman, and Figen Karadogan (2009), "What I Regret Most Are Lost Opportunities: A Theory of Regret Intensity," *Personality and Social Psychology Bulletin*, 35 (3), 385-97.

Berman, Jonathan Z., An T. K. Tran, John G. Lynch Jr, and Gal Zauberman (2016), "Expense Neglect in Forecasting Personal Finances," *Journal of Marketing Research*, 53 (4), 535—50.

Chernev, Alexander (2008), "The Role of Purchase Quantity in Assortment Choice: The Quantity-Matching Heuristic," *Journal of Marketing Research*, 45 (2), 171—81.

Cho, Eunice Kim, Uzma Khan, and Ravi Dhar (2013), "Comparing Apples to Apples or Apples to Oranges: The Role of Mental Representation in Choice Difficulty," *Journal of Marketing Research*, 50 (4), 505—16.

Clee, Mona A. and Robert A. Wicklund (1980), "Consumer Behavior and Psychological Reactance," *Journal of Consumer Research*, 6 (4), 389—405.

De Dreu, Carsten KW, Sander L. Koole, and Frans L. Oldersma (1990), "On the Seizing and Freezing of Negotiator Inferences: Need for Cognitive Closure Moderates the Use of Heuristics in Negotiation," *Personality and Social Psychology Bulletin*, 25 (3), 348—62.

Dhar, Ravi and Stephen M. Nowlis (2004), "To Buy or Not to Buy: Response Mode Effects on Consumer Choice," *Journal of Marketing Research*, 41 (4), 423—32.

Dhar, Ravi, Joel Huber, and Uzma Khan (2007), “The Shopping Momentum Effect,” *Journal of Marketing Research,* 44 (3), 370—78.

Disatnik, David and Yael Steinhart (2015), "Need for Cognitive Closure, Risk Aversion, Uncertainty Changes, and their Effects on Investment Decisions," *Journal of Marketing Research,* 52 (3), 349—59.

Fitzsimons, Gavan J. and Donald R. Lehmann (2004), “Reactance to Recommendations: When Unsolicited Advice Yields Contrary Responses,” *Marketing Science*, 23 (1), 82—94.

Ford, Thomas E. and Arie W. Kruglanski (1995), "Effects of Epistemic Motivations on the Use of Accessible Constructs in Social Judgment," *Personality and Social Psychology Bulletin,* 21 (9), 950—62.

Fox, Craig R., Rebecca K. Ratner, and Daniel S. Lieb (2005), "How Subjective Grouping of Options Influences Choice and Allocation: Diversification Bias and the Phenomenon of Partition Dependence," *Journal of Experimental Psychology: General,* 134 (4), 538—51.

Goldsmith, Kelly, Jing Xu, and Ravi Dhar (2016), “The Psychology of Basket Size: How Mental Representation and the Relationships between Products Affect Purchase Quantity,” *working paper.* Available at: https://ssrn.com/abstract=2802329.

Gollwitzer, Peter M. (1999), "Implementation Intentions: Strong Effects of Simple Plans," *American Psychologist,* 54 (7), 493—503.

Gollwitzer, Peter M. and Paschal Sheeran (2006), "Implementation Intentions and Goal Achievement: A Meta-Analysis of Effects and Processes," *Advances in Experimental Social Psychology,* 38, 69—119.

Grice, H. Paul (1975), “Logic and Conversation,” in *Syntax and Semantics: 3 Speech Acts*, ed. Peter Cole and Jerry L. Morgan, New York: Academic Press, 41–58.

Gu, Yangjie, Simona Botti, and David Faro (2013), "Turning the Page: The Impact of Choice Closure on Satisfaction," *Journal of Consumer Research,* 40 (2), 268—83.

Hilton, Denis J. (1995), "The Social Context of Reasoning: Conversational Inference and Rational Judgment," *Psychological Bulletin,* 118 (2), 248—71.

Ilyuk, Veronika and Lauren Block (2016),” The Effects of Single-Serve Packaging on Consumption Closure and Judgments of Product Efficacy,” *Journal of Consumer Research,* 42 (6), 858—78.

Inman, J. Jeffrey, Anil C. Peter, and Priya Raghubir (1997), "Framing the Deal: The Role of Restrictions in Accentuating Deal Value," *Journal of Consumer Research,* 24 (1), 68—79.

Johnson, Eric J., John Hershey, Jacqueline Meszaros, and Howard Kunreuther (1993), "Framing, Probability Distortions, and Insurance Decisions," *Journal of Risk and Uncertainty,* 7 (1), 35—51.

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica,* 47 (2), 263—92.

Kardes, Frank R., Maria L. Cronley, James J. Kellaris, and Steven S. Posavac (2004), "The Role of Selective Information Processing in Price-Quality Inference," *Journal of Consumer Research,* 31 (2), 368—74.

Kardes, Frank R., Bob M. Fennis, Edward R. Hirt, Zakary L. Tormala, and Brian Bullington (2007), "The Role of the Need for Cognitive Closure in the Effectiveness of the Disrupt-then-Reframe Influence Technique," *Journal of Consumer Research,* 34 (3), 377—85.

Kim, Hyeongmin (2013), "How Variety-Seeking Versus Inertial Tendency Influences the Effectiveness of Immediate Versus Delayed Promotions," *Journal of Marketing Research,* 50 (3), 416—26.

Kellas, Jody Koenig and Valerie Manusov (2003), "What's in a Story? The Relationship Between Narrative Completeness and Adjustment to Relationship Dissolution," *Journal of Social and Personal Relationships,* 20 (3), 285—307.

Kellas, Jody Koenig, Dawn Bean, Cherakah Cunningham, and Ka Yun Cheng (2008), "The Ex-Files: Trajectories, Turning Points, and Adjustment in the Development of Post-Dissolutional Relationships," *Journal of Social and Personal Relationships,* 25 (1), 23—50.

Kollat, David T. and Ronald P. Willett (1967), "Customer Impulse Purchasing Behavior," *Journal of Marketing Research,* 4 (1), 21—31.

Kruglanski, Arie W. and Donna M. Webster (1996), "Motivated Closing of the Mind: 'Seizing' and 'Freezing,'" *Psychological Review,* 103 (2), 263—83.

Kupor, Daniella M., Taly Reich, and Baba Shiv (2015), "Can't Finish What You Started? The Effect of Climactic Interruption on Behavior," *Journal of Consumer Psychology,* 25 (1), 113—19.

Lee, Kyoungmin and Sharon Shavitt (2009), "Can McDonald's Food Ever be Considered Healthful? Metacognitive Experiences Affect the Perceived Understanding of a Brand," *Journal of Marketing Research,* 46 (2), 222—33.

Lessne, Greg J. and Elaine M. Notarantonio (1988), "The Effect of Limits in Retail Advertisements: A Reactance Theory Perspective," *Psychology & Marketing,* 5 (1), 33—44.

Levav, Jonathan, Mark Heitmann, Andreas Herrmann, and Sheena S. Iyengar (2010), "Order in Product Customization Decisions: Evidence from Field Experiments," *Journal of Political Economy,* 118 (2), 274—99.

Levav, Jonathan, Nicholas Reinholtz, and Claire Lin (2012), "The Effect of Ordering Decisions by Choice-Set Size on Consumer Search," *Journal of Consumer Research,* 39 (3), 585—99.

Levin, Irwin P. and Gary J. Gaeth (1988), "How Consumers are Affected by the Framing of Attribute Information Before and After Consuming the Product," *Journal of Consumer Research*, 15 (3), 374—78.

Lieberman, Alicea, Kristen Duke, and On Amir (2019), "How Incentive Framing Can Harness the Power of Social Norms," *Organizational Behavior and Human Decision Processes,* 151, 118—31.

McKenzie, Craig R. M. and Jonathan D. Nelson (2003), "What a Speaker's Choice of Frame Reveals: Reference Points, Frame Selection, and Framing Effects," *Psychonomic Bulletin & Review,* 10 (3), 596—602.

McNeil, Barbara J., Stephen G. Pauker, Harold C. Sox Jr, and Amos Tversky (1982), "On the Elicitation of Preferences for Alternative Therapies," *New England Journal of Medicine,* 306 (21), 1259—62.

McShane, Blakeley B. and Ulf Böckenholt (2017), "Single-Paper Meta-Analysis: Benefits for Study Summary, Theory Testing, and Replicability," *Journal of Consumer Research,* 43 (6), 1048—63.

Miller, John H. (1978), "The Problematic of Endings in Narrative," in *Narrative Endings: A Special Issue of Nineteenth-Century Fiction,* ed. Alexander Welsh, Regents of the University of California: Publisher, 3—7.

Morwitz, Vicki G. and Gavan J. Fitzsimons (2004), "The Mere-Measurement Effect: Why Does Measuring Intentions Change Actual Behavior?," *Journal of Consumer Psychology,* 14 (1-2), 64—74.

Nickerson, David W. and Todd Rogers (2010), "Do You Have a Voting Plan? Implementation Intentions, Voter Turnout, and Organic Plan Making," *Psychological Science,* 21 (2), 194—99.

Nowlis, Stephen M., Ravi Dhar, and Itamar Simonson (2010), "The Effect of Decision Order on Purchase Quantity Decisions," *Journal of Marketing Research,* 47 (4), 725—37.

Park, C. Whan, Sung Youl Jun, and Deborah J. MacInnis (2000), "Choosing What I Want Versus Rejecting What I Do Not Want: An Application of Decision Framing to Product Option Choice Decisions," *Journal of Marketing Research,* 37 (2), 187—202.

Parker, Jeffrey R., and Rom Y. Schrift (2011), "Rejectable Choice Sets: How Seemingly Irrelevant No-Choice Options Affect Consumer Decision Processes," *Journal of Marketing Research,* 48 (5), 840—54.

Puto, Christopher P. (1987), "The Framing of Buying Decisions," *Journal of Consumer Research,* 14 (3), 301—15.

Prelec, Drazen, Birger Wernerfelt, and Florian Zettelmeyer (1997), "The Role of Inference in Context Effects: Inferring What You Want from What is Available," *Journal of Consumer Research,* 24 (1), 118—25.

Richter, David H. "Fable's End: Completeness and Closure in Rhetorical Fiction," *Modern Philology: Critical and Historical Studies in Literature, Medieval through Contemporary,* 76 (1), 106—108.

Roets, Arne and Alain Van Hiel (2011), "Item Selection and Validation of a Brief, 15-Item Version of the Need for Closure Scale," *Personality and Individual Differences,* 50 (1), 90—94.

Schrift, Rom Y., Jeffrey R. Parker, Gal Zauberman, and Shalena Srna (2017), "Multi-Stage Decision Processes: The Impact of Attribute Order on How Consumers Mentally Represent Their Choice," *Journal of Consumer Research,* forthcoming.

Schwarz, Norbert (1999), "Self-Reports: How the Questions Shape the Answers," *American Psychologist,* 54 (2), 93—105.

Schwarz, Norbert, Hans-J. Hippler, Brigitte Deutsch, and Fritz Strack (1985), "Response Scales: Effects of Category Range on Reported Behavior and Comparative Judgments," *The Public Opinion Quarterly,* 49 (3), 388—95.

Simonson, Itamar (1990), "The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior," *Journal of Marketing Research,* 27 (2), 150—62.

Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch Jr, and Gary H. McClelland (2013), "Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression," *Journal of Marketing Research*, 50 (2), 277—88.

Stilley, Karen M., J. Jeffrey Inman, and Kirk L. Wakefield (2010), "Planning to Make Unplanned Purchases? The Role of In-Store Slack in Budget Deviation," *Journal of Consumer Research*, 37 (2), 264—78.

Tannenbaum, David, Fox, Craig R., and Noah J. Goldstein (2013), "Partitioning Menu Options to Nudge Single-Item Choice," *working paper*. Available at: https://davetannenbaum.github.io/documents/pdepend.pdf.

Trope, Yaacov and Nira Liberman (2003), "Temporal Construal," *Psychological Review*, 110 (3), 403—21.

Tversky, Amos and Daniel Kahneman (1981), "The Framing of Decisions and the Psychology of Choice," *Science*, 211 (4481), 453—58.

---------- (1986), "Rational Choice and the Framing of Decisions," *Journal of Business*, 59 (4), S251-S278.

von Neumann, J., & Morgenstern, O. (1947), *Theory of Games and Economic Behavior, 2nd Ed*. Princeton, NJ: Princeton University Press.

Wansink, Brian, Robert J. Kent, and Stephen J. Hoch (1998), "An Anchoring and Adjustment Model of Purchase Quantity Decisions," *Journal of Marketing Research*, 35 (1), 71—81.

Webster, Donna M. and Arie W. Kruglanski (1994), "Individual Differences in Need for Cognitive Closure," *Journal of Personality and Social Psychology*, 67 (6), 1049—62.

Wernerfelt, Birger (1994), "Selling Formats for Search Goods," *Marketing Science*, 13 (3), 298—309.

Wohl, Jessica (2018), "Domino's Unseats Pizza Hut as Biggest Pizza Chain," *Ad Age*, February 20.

Xu, Alison Jing and Robert S. Wyer Jr. (2007), "The Effect of Mind-Sets on Consumer Decision Strategies," *Journal of Consumer Research*, 34 (4), 556—66.

Chapter 3.

HOW INCENTIVE FRAMING CAN HARNESS THE POWER OF SOCIAL NORMS

Alicea Lieberman, Kristen Elizabeth Duke, On Amir

Rady School of Management, University of California, San Diego, La Jolla, CA, 92093, USA

ABSTRACT

Incentives are an increasingly common tool used by organizations, managers, and policymakers to change behavior. We propose that more than just motivating behavior for monetary reasons, incentives also have an important, undiscovered consequence: they leak information about social norms. Four experiments reveal that framing an incentive as a surcharge, as compared to a discount, signals that the incentivized behavior is both more socially approved and more common. These implied norms lead individuals to experience emotions consistent with a desire to conform, motivating them to perform the incentivized behavior. Moreover, by shifting social norms, we find that incentives can influence behavior not only in the moment, but also downstream when there is no longer an active incentive. Further, merely being exposed to a surcharge (vs. discount) incentive—even without being financially affected by it—can increase performance of the behavior. These findings offer a novel perspective on the consequences of different incentive frames, while contributing to both organizational policy and practice by expanding the social norms messaging toolkit.

INTRODUCTION

Social norms influence behavior in powerful ways and can serve as a tool for large-scale behavior change. Indeed, perceptions of what is desirable or typical in a given context (Miller & Prentice, 1996) can drive behavior change at both the individual level (e.g., increasing recycling; Cialdini, Reno, & Kallgren, 1990), and the organizational level (e.g., improving group cohesion and performance; George & Jones, 1997, Stewart et al., 2012; De Jong, Bijlsma-Frankema, & Cardinal, 2014), as individuals are deeply motivated to gain and maintain acceptance from their peers (Barker 1993; Cialdini, Kallgren, & Reno, 1991; Perkins & Berkowitz, 1986). Social norms often evolve organically over time as individuals learn what is approved of and what is common in a community (Miller & Prentice, 1996; Paluck & Shepherd, 2012; Feldman, 1984; Ehrhart & Naumann, 2004). They can also be influenced through multiple channels, such as the personality characteristics (Gonzalez-Mule, 2014) or shared interactions of a group (Bettenhausen & Murnighan, 1985; Hackman, 1992). Commonly, efforts to shift social norms involve explicit messaging campaigns that state which behaviors are most normative. Such campaigns have successfully altered a wide range of behaviors, including decreasing binge drinking (Haines & Spear, 1996), influencing healthy food choices (Robinson et al., 2014), and increasing hotel towel re-use (Goldstein, Cialdini, & Griskevicius, 2008). However, such direct approaches also come with other risks, as advertising can irritate people (Aaker & Bruzzone, 1985) or even cause them to avoid an ad altogether (Speck & Elliot, 1997). In this paper, we investigate whether social norms can be communicated indirectly using a simple tool commonly used by organizations, managers, and policymakers: small

monetary incentives. In particular, we propose that the framing of an incentive can influence people's beliefs about how approved of and how common the incentivized behavior is, leading to important behavioral consequences.

Monetary incentives can be framed in a negative or positive light: either as additional costs one must pay for behaving in an undesired manner (i.e., a surcharge) or as costs deducted for engaging in a desired manner (i.e., a discount). For example, companies may charge higher premiums on health insurance plans for employees with unhealthy lifestyle behaviors such as smoking (Abelson, 2011; CMS, 2013), or may offer their employees discounts on insurance premiums for healthy lifestyle behaviors such as participating in wellness programs (kff, 2016). Customers may also pay surcharges or earn discounts when shopping, as when paying 5-cent charges for using plastic bags (National Conference of State Legislatures, 2017) or earning 5-cent discounts for bringing their own reusable bags (Target, 2017). Surcharge and discount incentives are increasingly common tools for behavior change, predicated on the idea that incentives are motivating for their monetary value (DellaVigna & Pope, 2017). However, we propose that incentives also have an important signaling value. Specifically, we demonstrate that people who encounter an incentive framed as a surcharge infer that the targeted behavior is both more approved of and more common than when the incentive is framed as a discount. The prospect of violating this perceived norm elicits embarrassment and guilt, two negative, self-conscious emotions associated with the desire to conform to perceived norms. In turn, these anticipated emotional reactions drive individuals to carry out the encouraged behavior. Thus, the current research proposes social norm leakage as a novel consequence of

incentive framing, and further, proposes that this leakage plays a key role in the power of incentives to change behavior.

In addition to eliciting norm-related emotions and changing behavior in the moment, this framework also implies two additional novel predictions. Namely, because people strive to emulate social norms, exposure to an incentive can influence their behavior not only initially, but also downstream when there is no longer an active incentive. Furthermore, because of this norm signaling, merely learning about the existence of an incentive (as when seeing it written on a sign) can affect whether individuals carry out the behavior, even when the incentives do not financially impact them directly. In this manner, our research contributes to the organizational behavior and policy literatures by offering a deeper understanding of how and why incentives affect behavior, and by providing novel insights on how to effectively frame such policies to capitalize on the power of social norms.

Information Leakage from Policies

Past work suggests that the decision to enact certain policies may provide insight into which behaviors are prescribed within a given community (Lapinski & Rimal, 2005). In a similar way, we propose that in addition to the enactment of an incentive policy revealing information about organizational or community expectations, how that incentive is framed also conveys important social norm information that can influence people's

perceptions, beliefs, and actions. We find support for this notion in the information leakage literature.

Information leakage is the phenomenon in which the structure of a policy, choice set, or other feature of the environment signals information to individuals (McKenzie, 2004; Sher & McKenzie, 2006), and has been used to explain certain framing effects. Consider the default effect, in which people tend to stick with the default (or "do nothing") option of a choice set (Johnson & Goldstein, 2003). An information leakage account proposes that this choice arises, in part, because individuals interpret an organization's selection of a default as an implicit recommendation—that is, people perceive that the institution implementing the policy recommends that people perform the default action (McKenzie, Liersch, & Finkelstein, 2006). In this way, the structuring of a choice set can provide a subtle cue for decision-makers about how the choice architect expects them to behave (e.g., Krijnen, Tannenbaum, & Fox, 2018)–the "right" choice. The framing of a policy may also leak information about a company or policymaker's attitudes toward those affected. For example, Tannenbaum and colleagues (2013) demonstrate that disincentive policies (e.g., increased premiums for being overweight), can lead people to infer that the company imposing the policy holds negative attitudes toward the targeted individuals (e.g., overweight employees), while positive incentives (e.g., decreased premiums) do not lead to such inferences.

In a similar manner, we suggest that the framing of an incentive can influence people's perceptions of the expected or "right" behavior. In particular, we propose that incentive framing can affect perceptions of injunctive norms: what people perceive to be

the approved behavior—that is, what one "ought" to do (Cialdini, Reno, & Kallgren, 1990, 1991). In other words, incentive framing can suggest to individuals what behaviors are approved of and accepted in a given context, projecting an expectation of compliance (Fehr & Schurtenberger, 2018). This idea follows from work on policy framing. Policymakers can frame policies to either encourage a desirable behavior or to discourage an undesirable behavior, a choice that can influence the likelihood that the policy will be accepted by its stakeholders (Evers et al., 2016). Disadvantaging policies (e.g., punishments imposed on individuals who act in an undesired way) tend to be more accepted when the incentivized behavior is considered obligatory, such as required community service (Evers et al., 2016). In contrast, advantaging policies (e.g., rewards offered for acting in a desired way) tend to be more accepted when the incentivized behavior is viewed as optional, such as voluntary community service (Evers et al., 2016). This idea converges with related notions from the morality and economics literatures. Specifically, using social dilemma games, Mulder (2008) finds that punishments (vs. rewards) can signal that a behavior is morally obligatory (vs. voluntary). Similarly, Fehr and Fishbacher (2004) find that people punish others who violate cooperation norms, suggesting that sanctions may serve as a form of norm enforcement. Together, these findings suggest that the framing of a policy may signal something about the social expectations of the targeted behavior.

In our framework, surcharges are much like disadvantaging policies: they are penalties imposed on individuals who act in an undesired way, and thus should be more accepted when the behavior they incentivize is perceived as relatively obligatory. Organizations are more likely to enact policies supported by their constituents, and prior

research demonstrates that messages are the most persuasive when designed to reflect normative ideations already accepted by the group (Payne, 2001). Analogously, we propose that the most effective incentives will be those designed to resonate with their target audience by reflecting ideas already accepted by the community. Accordingly, we suggest that the presence of a surcharge signals that the targeted behavior is more of an obligation, an action one ought to do (i.e., an injunctive norm). In contrast, discounts are advantaging policies that benefit individuals who act in a desired way. Thus, they should be more accepted when they incentivize a behavior that is perceived as relatively voluntary, signaling that the behavior is not necessarily a social "ought." In this way, the framing of an incentive may leak that the behavior is more of an injunctive norm.

Note that it is possible for individuals to construe surcharges as a form of punishment, and discounts as a form of reward. However, in our framework, surcharges are simply additional costs, and discounts are simply cost deductions. Thus, while surcharges may be viewed as a form of punishment relative to discounts, we suggest that surcharges will signal that the targeted behavior is more normative regardless of whether or not people construe them as explicit punishments, reducing potential concerns of reactance or anger (e.g., Balliet, Mulder, & Van Lange, 2011). Employees, customers, and other individuals may not necessarily perceive surcharges as punishments (and discounts as rewards)—they may instead perceive them as cost passing tactics (i.e., that the organization or retailer is passing on the cost of supplying additional products or services to the employee or customer) or, either explicitly or implicitly, as merely small incentives designed to nudge behaviors. Thus, while construing a surcharge as a punishment would

likely only strengthen the signal that the incentivized behavior is a social a norm, our theory suggests that this signal will remain even if people do not view the incentive in this light. Indeed, we demonstrate that surcharges (vs. discounts) still project stronger social norms and serve as stronger drivers of behavior even when adjusting for the extent to which individuals perceive these incentives as punishments and rewards.

We further propose that the framing of an incentive can signal not only what behavior is most approved of, but also what behavior is most common—that is, the descriptive norm (Cialdini, Reno, & Kallgren, 1990, 1991). Often, people's perceptions of the approved behavior are consistent with how most people tend to behave—that is, injunctive and descriptive norms often (although do not always) converge (Erikkson, Strimling, & Coultas, 2015; Lapinski & Rimal, 2005). Thus, framing an incentive as a surcharge (vs. a discount) may signal not only that performing the incentivized behavior is more approved of and more expected (an injunctive norm), but also that it is more typical (a descriptive norm).

In addition to the common linkage between injunctive and descriptive norms, a rational account would also suggest that surcharges would likely only be successfully enacted in contexts where the incentivized behavior is already common, and thus few people would incur the extra charge. To illustrate our logic, consider a coffee shop that imposes a 10-cent surcharge on customers who do not bring their own mugs. If bringing a mug were uncommon, then many customers would be required to pay this additional charge. Because customers are motivated to avoid paying an extra cost, even if it is small, implementing a surcharge when the incentivized behavior is uncommon may result in

customers choosing to go to another shop, making it unlikely that a store would enact, or maintain, such a policy. Thus, the existence of a surcharge may suggest that few people behave in the undesired way (i.e., few people must pay the charge), signaling that the desired (i.e., incentivized) behavior is more common—a stronger descriptive norm.[11] On the other hand, consider a coffee shop that offers a 10-cent discount to customers who bring their own mugs. It is highly unlikely that a shop owner would offer this policy when the behavior is already common, as it would unnecessarily pay people for a behavior they are already performing.[12] Said another way, because people are more likely to accept policies that are aligned with their perceptions of the targeted behavior (Evers et al. 2016), a shop owner is unlikely to enact a policy that is misaligned with the community's beliefs, and further, a misaligned policy is unlikely to stay enacted. Thus, we predict our proposed process should apply equally at both the enactment and maintenance stages of policy implementation. Thus, together, our theorizing suggests that surcharges (vs. discounts) leak that the incentivized behavior is more approved of and more common. Next, we draw

[11] Note that a related, but alternative, logic would also make the same prediction. Specifically, individuals may infer that there is a standard, baseline outcome to which adjustments can be made (and they may assume that it is more efficient to set baselines in a way that adjustments are rare). In this case, surcharges and discounts would serve as adjustments that arise when behavior does not align with the baseline outcome. As such, the presence of a surcharge could suggest that needing to pay this charge (i.e., not performing the desired behavior) deviates from the standard baseline behavior and, thus, is rare. Similarly, the presence of a discount could suggest that performing the desired behavior (and earning this benefit) deviates from the standard baseline behavior and, thus, is rare. This inference process can explain why individuals infer that performing the desired behavior (and avoiding the charge/earning the discount) is more common under a surcharge than a discount. We thank an anonymous reviewer for this suggestion.

[12] A related practical argument would suggest that if it is effortful for a shop owner to collect a surcharge, it is less practical to impose this policy when many customers would be charged (i.e., when the desired behavior is rare), while it is similarly impractical to offer a discount when many customers would earn it (i.e., when the desired behavior is common). This logic suggests another pathway through which incentive framing could signal descriptive norms surrounding the incentivized behavior. We thank the review team for these insights.

on social norms research to propose that these inferences carry important consequences for how individuals behave.

Social Norms and Their Consequences

As social animals, people's behaviors are strongly shaped by their perceptions of social norms (Ajzen & Fishbein, 1980; Asch, 1951; Fishbein & Ajzen, 1975), both how one ought to and is expected to behave (injunctive norms) and how most others behave (descriptive norms; Cialdini & Trost, 1998; Cialdini, Reno, & Kallgren, 1990, 1991; Goldstein, Cialdini, & Griskevicius, 2008). While injunctive and descriptive norms are often congruent (Lapinski & Rimal, 2005), they are conceptually distinct, and both types of norms provide important information about how to act in a given setting (e.g., Cialdini, Kallgren, & Reno, 1991). Thus, while we primarily focus on injunctive norms throughout this paper, we also measure and report descriptive norms. Notably, research suggests that it is perceptions of the norm, more so than the actual norm, that influence behavior (Berkowitz, 2004). Accordingly, many social norm interventions aimed at changing a community norm do so by attempting to shift the subjective perceptions of the norm (Tankard & Paluck, 2016), which can change behaviors even when the behavior is not performed by the majority (Goldstein, Cialdini, & Griskevicius, 2008). Thus, throughout the paper when discussing a "norm," we are alluding to this subjective perception of the norm.

Given the well-established power of social norms, they are featured as a key construct in many behavior change models (e.g., Ajzen, 1991; Fishbein & Yzer, 2003). This provides reason to suspect that if surcharges leak stronger social norms than discounts, they will also lead to greater behavior changes. Drawing on individuals' fundamental desire to fit in with their peers, we propose that by shifting perceptions that a behavior is more normative, surcharges (vs. discounts) will also better motivate behavior. Moreover, this increased motivation should be strongest among individuals who have the highest tendency to conform to social norms (Bearden & Rose, 1990). That is, we predict that the superior effectiveness of surcharges over discounts will be moderated by the extent to which individuals tend to conform to their peers.

One key mechanism by which social norms influence behavior is through anticipated emotions. Past research demonstrates that violating a perceived social norm can elicit self-conscious emotions (e.g., Hareli & Parkinson, 2008; Keltner & Anderson, 2000), which include embarrassment, guilt, shame, and pride. Self-conscious emotions are unique in so far as they all involve a heightened awareness and evaluation of the self (e.g., Tangney & Fischer, 1995), resulting from appraisals of one's own actions. As a result, self-conscious emotions help individuals maintain or improve their social status and avoid group rejection (Keltner & Buswell, 1997; Tracy & Robins, 2004).

In the current research, we focus on the two self-conscious emotions that are action-oriented and motivate social behavior, and thus most relevant to the contexts we investigate: embarrassment and guilt (see Appendix for a discussion of shame and pride, two other self-conscious emotions that share some conceptual features with embarrassment

and guilt). Embarrassment often arises when violating a perceived social convention, typically in the presence of an external audience (Keltner & Buswell, 1997), and leads to concern that a social mishap may threaten the individual's social identity or position within a social group. Accordingly, embarrassment motivates appeasement behaviors designed to elicit reconciliation with others and signal a commitment to upholding social norms (Keltner, Young, & Buswell, 1997).

Guilt involves an appraisal of personal responsibility for a specific action or choice (Lewis 1971; Tracy & Robbins, 2004; Duke & Amir, 2018), and results when individuals feel that their behavior falls short of appropriate, desired conduct (Baumeister, Stillwell, & Heatherton, 1994). Like embarrassment, guilt also arises when people perceive themselves as violating social conventions (Keltner & Buswell, 1996; Tangney & Dearing, 2002; Fehr & Schurtenberger, 2018), and this emotion motivates atonement behaviors aimed at amending the social harm imposed (e.g., Lindsay-Hartz, De Rivera, & Mascolo, 1995).

Thus, we propose that because surcharges (vs. discounts) signal that the incentivized behavior is more desired, expected, and common to perform, the prospect of failing to uphold this norm will lead to the anticipation of embarrassment and guilt. In order to prevent these negative feelings and their associated consequences, individuals will be more motivated to perform the behavior. Said another way, anticipated embarrassment and guilt may serve as a key pathway through which incentive framing affects behavior. These self-conscious emotions of embarrassment and guilt have been intimately linked to social norms violations; accordingly, in the remainder of this paper, we refer to them as "norm-related emotions."

Critically, a key consequence of this social norm mechanism is the possibility for behavior change to linger and carry over to later instances where incentives are not in place, but where the same social norms are expected. Indeed, past work has documented that efforts to shift social norms can have long-lasting effects (Neighbors, Larimer, & Lewis, 2004; Schultz et al., 2007). In this research we posit that because surcharges leak information about social norms, they hold the power to have an extended impact on individuals' perceptions and behaviors. We further propose that it is the presence of a surcharge, even when it is not experienced directly, that signals that a behavior is more normative—affecting anyone who is simply exposed to the incentive. In other words, experiencing or merely observing an incentive at one point in time can shift perceptions of social norms and thus influence behaviors after the incentive is no longer in place, as well as in other similar locations without active incentives. In this way, everyday incentives used by employers, policymakers, and marketers may have a surprising and previously undiscovered impact on individuals' future actions.

A Note on Loss Aversion

Social norms are not the only mechanism through which incentives can change behavior; individuals are undoubtedly also motivated by incentives for their pure monetary value. Further, the framing of an incentive can influence how individuals react to and value it. Prior research suggests that losses are more impactful and perceived as larger than equivalent gains, an asymmetry termed "loss aversion" (Kahneman, Knetsch, & Thaler,

1991; Kahneman & Tversky, 1979). In some situations, people may evaluate the presence of an incentive in comparison to a situation with no incentive at all. In this case, they may perceive the payment of a surcharge as a potential loss, and the receipt of a discount as a potential gain (Kahneman & Tversky, 1986; Thaler, 1980). With this idea in mind, some research uses loss aversion to describe why individuals might react more strongly to surcharges than to discounts (Nasiry & Popescu, 2011; Homonoff, 2015; Poortinga, 2017). However, attributing surcharges' motivational power entirely to loss aversion requires that individuals value losses to be substantially higher than gains. For example, Homonoff (2015) evaluates surcharge and discount incentives in the field and calculates a loss aversion coefficient between 5 and 14, which is quite high relative to other observed measures (~2; Abdellaoui, Bleichrodt, & Paraschiv, 2005; Camerer, 2005). This suggests that another mechanism might be at play—which we propose is the leakage of social norms. In this research, we also offer predictions and demonstrate evidence for several important consequences that would not be predicted by valuation of the incentives (as a result of perceiving them as losses or gains; i.e., loss aversion), and instead would only arise if incentive framing leaked social norms. We discuss these differences in each study and return to the psychology of loss aversion in the General Discussion.

OVERVIEW OF STUDIES

Four studies test our proposed account. Study 1 demonstrates that surcharges (vs. discounts) heighten expectations to conform to an environmentally friendly behavior and

increase the perceived prevalence of that behavior—thus leaking both an injunctive and descriptive norm. Further, it finds that failing to conform under a surcharge (vs. discount) elicits stronger norm-related emotions which, in turn, increase one's intention to perform the incentivized behavior. Study 2 conceptually replicates these findings in a new context and provides further process evidence of a social norms account by demonstrating that exposure to a surcharge at one location influences intentions to engage in the focal behavior in another similar location where no incentive is offered. Study 3 provides additional evidence that the social information leaked by surcharges is a key driver of their ability to motivate behavior by demonstrating that one's likelihood of engaging in a behavior incentivized with a surcharge (vs. a discount) is moderated by individual differences in sensitivity to social influence. That is, the motivational power of surcharges over discounts is driven by those individuals who care the most about conforming to the expectations of their peers. Finally, Study 4 provides behavioral evidence supporting the notion of carryover effects. Specifically, it shows that exposure to a surcharge (vs. a discount) at one point in time increases incidence of the encouraged behavior at a second point in time when no incentive is offered. Moreover, we find that this effect holds even among individuals who merely witnessed the incentive at the earlier point in time and were not financially affected by it. In all studies, we predetermined sample size and report all data exclusions (if any), all manipulations, and all measures.

STUDY 1: SURCHARGES LEAK STRONGER SOCIAL NORMS AND INCREASE BEHAVIORAL INTENTIONS

Study 1 investigated the decision to bring a reusable mug to a coffee shop, a behavior incentivized by either a surcharge or a discount. This context mirrors a growing real-world phenomenon: organizations increasingly offer surcharge or discount incentives to encourage reuse behaviors. For example, in early 2018, several UK Starbucks shops launched a trial "latte levy," requiring customers to pay a surcharge if they did not bring a reusable mug (The Guardian, 2018). In contrast, US Starbucks shops offer a 10-cent discount for customers who bring reusable mugs (Starbucks, 2017). Study 1 examined whether framing the incentive as a surcharge (vs. a discount) signals that bringing a reusable mug is more of a social norm which, in turn, increases people's anticipated self-conscious emotions and intention to perform the incentivized behavior.

Method

Three hundred two Mechanical Turk workers (43.4% female, M_{age} = 34.6 years, SD_{age} = 11.9 years) participated. Participants were randomly assigned to one of two conditions (incentive framing: surcharge vs. discount) in a between-subjects design. All participants read, "Imagine that you have just moved to a new town, Newbury, and are looking for opportunities to get to know your new neighborhood and your new neighbors! You see a flyer for your local coffee shop and decide to go check it out!" On the next page, participants saw a flyer for "Newbury Bean House," with text varying by condition (see Appendix for stimuli). For those in the discount condition, the sign read, "Customers will

get a 10-cent discount for bringing their own mugs!" while those in the surcharge condition read, "Customers will be charged 10 cents for not bringing their own mugs!"

Participants then responded to two items in randomized order, capturing their perceptions of the injunctive norm: "Most people at this shop think that everyone ought to bring their own reusable coffee mug" and "Most people at this shop think that I should bring my own reusable coffee mug" (1 = No, definitely not to 7 = Yes, definitely). These items were adapted from prior research investigating perceived injunctive norms (White et al., 2009; Smith-McLallen & Fishbein, 2008), and were averaged to form a single scale (r = .83). Participants then provided their perceptions of the descriptive norm, answering: "Out of every 100 customers, how many do you think bring their own reusable mug?" (0-100).

On the following page, participants read, "Take a moment to think about how you would feel if you did NOT bring your own coffee mug to this shop." They then responded to 10 items capturing their anticipated guilt and embarrassment, in randomized order (1 = Not at all to 7 = Very much). Five items ($\alpha = .95$) were drawn from the guilt subscale from the State Shame and Guilt Scale (Marschall, Sanftner, & Tangney, 1994) and slightly altered to capture anticipated (vs. in-the-moment) emotions. These items were, "I would feel remorse, regret"; "I would feel tension about something I have done"; "I would not be able to stop thinking about something bad I have done"; "I would feel like apologizing, confessing"; and "I would feel bad about something I have done." Five items ($\alpha = .94$) were designed to capture anticipated embarrassment, based on Keltner and Buswell (1997)'s description of the characteristics and psychological experience of embarrassment.

These items were, “I would feel awkward”; “I would feel foolish”; “I would feel self-conscious”; “I would feel nervous”; and “I would feel worried.”

On the next page, participants reported their intention to conform to the incentivized behavior, answering: “How likely would you be to bring your own reusable mug to this shop?” (1 = Not at all to 7 = Very). Finally, participants provided demographic information.

Results

Injunctive Norms. In line with our predictions, participants perceived significantly higher injunctive norms regarding reusable mug use in the surcharge condition ($M = 5.73$, $SD = 1.26$) than in the discount condition ($M = 5.01$, $SD = 1.44$; $t(300) = 4.65$, $p < .001$, 95% $CI_{difference} = [.42, 1.03]$, $d = .31$).

Descriptive Norms. Participants also expected that significantly more customers would bring a reusable mug in the surcharge condition ($M = 60.66$, $SD = 21.72$) than in the discount condition ($M = 49.53$, $SD = 23.40$; $t(300) = 4.28$, $p < .001$, 95% $CI_{difference} = [6.02, 16.25]$, $d = .49$), indicating a stronger descriptive norm.

Norm-Related Emotions. As predicted, participants who read about a surcharge (vs. discount) also anticipated feeling guiltier (surcharge: $M = 3.20$, $SD = 1.71$ vs. discount: $M = 2.66$, $SD = 1.75$; $t(300) = 2.70$, $p = .007$, 95% $CI_{difference} = [.15, .93]$, $d = .31$) and more embarrassed (surcharge: $M = 3.63$, $SD = 1.75$ vs. discount: $M = 2.76$, $SD = 1.75$; $t(300) = 4.32$, $p < .001$, 95% $CI_{difference} = [.47, 1.26]$, $d = .50$) for failing to bring their own mug. The

guilt and embarrassment scales were highly correlated (r = .92) and were accordingly combined to form a single scale for the serial mediation analysis below.

Intention. Importantly, participants indicated significantly higher intention to bring their own mug in the surcharge condition (M = 5.44, SD = 1.45) than in the discount condition (M = 4.80, SD = 1.68; t(300) = 3.57, p < .001, 95% $CI_{difference}$ = [.29, 1.00], d = .41).

Serial Mediation. To test whether the higher intention in the surcharge (vs. discount) condition was driven by participants' norm perceptions and anticipated emotions, we conducted a serial mediation analysis with 5,000 bootstrapped samples (Hayes, 2013, model 6). Incentive framing (surcharge vs. discount) was the predictor, perceived injunctive norms was the first mediator, norm-related emotions served as the second mediator, and intention to bring a reusable mug was the dependent variable. The serial mediation confirmed the predicted path: the surcharge boosted norm perceptions, heightening norm-related emotions which, in turn, increased intention (serial paths $a_1 \times d_{21} \times b_2$ = .04, SE = .02, 95% CI: [.02, .08]; see Appendix for model). We also find significant patterns when replacing the injunctive norms measure with the descriptive norms measure (see Appendix).

Discussion

Study 1 demonstrates that surcharges (vs. discounts) lead individuals to perceive that conforming to an incentivized behavior is both more of an injunctive and descriptive

norm. Consistent with our predictions, and with prior research that norm violations elicit embarrassment and guilt (e.g., Hareli & Parkinson, 2008), these norm perceptions increased anticipated embarrassment and guilt for failing to carry out the incentivized behavior. This, in turn, increased participants' intention to bring a reusable mug under a surcharge (vs. under a discount).

The Appendix presents a conceptual replication of these results and addresses one alternative explanation by changing the description of the incentives. In Study 1, the surcharge was imposed on participants for not performing a behavior (i.e., not bringing a reusable mug), while the discount was offered for performing a behavior (i.e., bringing a reusable mug). In the experiment presented in the Appendix, both incentives are designed such that the highlighted behavior involves performing a behavior (i.e., "Customers will get a 10 cent discount for bringing their own bags" and "Customers will be charged 10 cents for using our bags"). Even in this case, participants still perceived stronger injunctive and descriptive norms and anticipated stronger norm-related emotional reactions under surcharges relative to discounts.

Importantly, our theory suggests that surcharges (vs. discounts) more strongly influence behaviors because they leak social norms. Thus, we predict that incentive framing should influence behaviors not only in the location offering the financial incentive, but also in other locations that do not offer any incentives, but where a similar norm might be expected. We test this conjecture in Study 2.

STUDY 2: SURCHARGES INCREASE DOWNSTREAM BEHAVIORAL INTENTIONS

Study 2 replicated the effect of incentive framing on social norm perceptions and anticipated norm-related emotions in a different, increasingly common, context. To encourage reusable bag use, surcharges and discounts are being implemented by organizations (Target, 2017) as well as state and local policymakers (National Conference of State Legislatures, 2017). A recent archival study investigating the effectiveness of such policies found that while a bag discount had little impact on behavior, a bag surcharge tripled reusable bag use—an effect the researcher attributed primarily to loss aversion (Homonoff, 2015).[13] Study 2 provides deeper insight into this finding, presenting evidence that leaked social norms may contribute to this greater effectiveness of surcharges over discounts.

Further, Study 2 investigated whether the social norms signaled via surcharges at one location would carry over to another nearby location with the same reference group, but where no incentives were offered. We predicted that a surcharge at one store would signal the neighborhood norm of bringing reusable bags, and thus bringing bags would also be perceived as the norm at another local store frequented by the same community members—even though there was no financial incentive to bring bags to this other

[13] Note that Homonoff (2015) also discusses the possibility of the enactment of government-imposed taxes or surcharges signaling a shift in social norms. She collected measures pre- and post-policy implementation and found no significant differences. However, most of the measures categorized as social norms in that work diverge from the definitions of injunctive and descriptive norms used in the present research. The items that were more aligned with our definition were directionally consistent with our theorizing.

location. As a result, we predicted that participants exposed to a surcharge (vs. discount) at one location would be more inclined to bring their own reusable bags to the second store where there was no financial incentive.

Method

Six hundred two Mechanical Turk workers (47.2% female, M_{age} = 35.8 years, SD_{age} = 12.0 years) participated. Participants were randomly assigned to one of two conditions (incentive framing: surcharge vs. discount) in a between-subjects design. All participants read, "Imagine that you have just moved to a new town, Newbury, and are looking for opportunities to get to know your new neighborhood and your new neighbors! You need groceries for the week and decide to go grocery shopping! Your new neighborhood has two local stores that are both very popular among your neighbors. You head to one of these neighborhood stores to do your shopping for the week." On the next page, participants were told that as they entered the neighborhood store they saw a sign. The sign presented the store's bag policy with text varying by condition. In the discount condition the sign indicated, "Customers will get a 10-cent discount for bringing their own bags!" while the sign in the surcharge condition read, "Customers will be charged 10 cents for not bringing their own bags!" (see Appendix for stimuli).

Participants then responded to the two injunctive norm measures from Study 1 (r = .85), in randomized order: "Most people at this shop think that everyone ought to bring their own reusable bags" and "Most people at this shop think that I should bring my own

reusable bags" (1 = No, definitely not to 7 = Yes, definitely). They then responded to the descriptive norm measure from Study 1: "Out of every 100 customers, how many do you think bring their own reusable bags?" (0-100). On the following page, participants reported their anticipated guilt and embarrassment in the same manner as Study 1.

On the next page, participants read, "The following weekend you need to go grocery shopping for the upcoming week. You decide to check out the second neighborhood store. This store does not offer an incentive for bringing reusable bags but otherwise is similar to the first store, carries similar inventory, and most of your neighbors shop at both locations." They then reported their perceptions of "the norm" at this incentive-less store, answering, "What do you think "the norm" (standard) is for customers at this neighborhood store?" (to bring reusable bags; to not bring reusable bags). This measure was meant to capture whether inferred social norms would carry over from the first store to the second store. Thereafter, they reported their intention, answering, "How likely would you be to bring your own reusable bags to this neighborhood store?" (1 = Not at all to 7 = Very). Finally, participants provided demographic information.

Results

Injunctive Norms. In this context, participants again perceived significantly higher injunctive norms regarding reusable bag use in the surcharge condition ($M = 5.99$, $SD = 1.11$) than in the discount condition ($M = 5.22$, $SD = 1.48$; $t(600) = 7.26$, $p < .001$, 95% $CI_{difference} = [.57, .99]$, $d = .59$).

Descriptive Norms. Participants also expected that significantly more customers would bring reusable bags in the surcharge condition (M = 68.93, SD = 20.60) than in the discount condition (M = 52.67, SD = 23.42; $t(600) = 9.04$, $p < .001$, 95% $CI_{difference}$ = [12.72, 19.79], d = .74), indicating a stronger descriptive norm.

Norm-Related Emotions. Participants who read about a surcharge (vs. discount) again anticipated feeling guiltier (surcharge: M = 2.97, SD = 1.70 vs. discount: M = 2.60, SD = 1.64; $t(600) = 2.70$, $p = .007$, 95% $CI_{difference}$ = [.10, .63], d = .22) and more embarrassed (surcharge: M = 3.39, SD = 1.75 vs. discount: M = 2.86, SD = 1.67; $t(600) = 3.77$, $p < .001$, 95% $CI_{difference}$ = [.25, .80], d = .31) for failing to bring their own bags (see Figure 3.1).

Norm Perceptions at the Incentive-less Store. Even though the second neighborhood store had no incentive for reusable bag use, participants still perceived stronger norms of usage in the surcharge (vs. discount) condition, as predicted. Significantly more surcharge participants perceived that the "norm" was to bring bags (vs. not bring bags) to this store (47.83%) than did discount participants (32.34%; $\chi^2(1, N = 602) = 15.03$, $p < .001$, 95% CI [7.58%, 23.11%], $\varphi = .16$).

Intention at the Incentive-less Store. In line with our theorizing, participants who had learned about a surcharge at one store in town were significantly more likely to bring their own reusable bags to the second neighborhood store that did not offer an incentive (M = 4.43, SD = 2.00) than were those who had learned about a discount at the first location (M = 4.07, SD = 1.93; $t(600) = 2.26$, $p = .02$, 95% $CI_{difference}$ = [.05, .68], d = .18; see Figure 3.2).

Serial Mediation. To test whether the higher downstream intention in the surcharge (vs. discount) condition was driven by participants' norm perceptions and anticipated emotions at the first grocery store, we conducted a serial mediation analysis with 5,000 bootstrapped samples (Hayes, 2013, model 6). Incentive framing (surcharge vs. discount) was the predictor, perceived injunctive norms at store 1 was the first mediator, norm-related emotions at store 1 served as the second mediator (r = .88), and intention to bring reusable bags to store 2 was the dependent variable. The serial mediation confirmed the predicted causal path: the surcharge boosted norm perceptions at store 1, heightening norm-related emotions which, in turn, increased intention at store 2, where no incentive was offered (serial paths $a_1 \times d_{21} \times b_2 = .05$, SE = .02, 95% CI: [.03, .09]; see Figure 3.3). Consistent with Study 1, we find significant patterns when replacing the store 1 injunctive norms measure with the store 1 descriptive norms measure (see Appendix for model).

Discussion

Study 2 again finds that surcharges project stronger injunctive and descriptive norms than do discounts in a context that reflects a growing policy trend: the use of small financial incentives to encourage reusable bag use. Moreover, a broad implication of a social norms account is the potential for carry over effects. Study 2 demonstrated that participants who saw a surcharge at one neighborhood store perceived that bringing a bag was more socially normative at that store and anticipated feeling stronger norm-related emotions for failing to bring a bag there. This, in turn, led them to perceive reusable bags

as more normative at another local store where no incentive was in place, and increased their intention to bring a bag to this other incentive-less location. An important feature of this study is that both locations belonged to the same general community, so it would be reasonable to expect social norms to transfer between the locations. Thus, while it is not necessary for both locations to have the exact same clientele for our effects to hold, our theorizing suggests that carry over effects are most likely to arise when the locations, at the very least, belong to similar communities.

In the next study, we investigate a new behavior and provide further support for a social norms account by testing whether individuals' susceptibility to social influence moderates their intention to comply with surcharge (vs. discount) incentives.

STUDY 3: THE TENDENCY TO CONFORM TO NORMS MODERATES RESPONSES TO INCENTIVES

Study 3 had three key goals. First, to provide further support for a norms-based account, we tested whether intention to comply with a surcharge (vs. discount) is moderated by an individual's tendency to conform to the norms of their peer group. Prior research suggests that individuals vary in the extent to which they are sensitive to information about what others are doing (e.g., Lennox & Wolfe, 1984). This individual difference in sensitivity to perceived norms has been shown to moderate the influence of social pressure on individuals' tendency to conform (Bearden & Rose, 1990). We posited that surcharges would again lead to greater behavioral intention relative to discounts,

driven by the social norm information they leak. Thus, we predicted that the gap in intention to carry out the targeted behavior under a surcharge should be larger among individuals who have the strongest tendencies to conform to peer influence.

Second, we generalized to a new context and behavior. Specifically, we investigated whether our theory extends to situations where the incentivized behavior is not inherently moral. Some prior work suggests that punishments (vs. rewards) can signal whether morally driven behavior is obligatory and can influence judgments of an individual's moral character (Mulder, 2008, 2016). It is possible that individuals may construe surcharges as punishments, and this construal could lead them to draw inferences about the morality of the behavior. However, note that at their core, surcharges are simply costs imposed (while discounts are costs removed) for behaving in a certain manner. Also note that while injunctive norms may signal what "should be done" and thus carry a moral flavor, this is distinct from whether the behavior itself is perceived as moral or immoral. Thus, diverging from prior research, in our framework it is not necessary to perceive the behavior as moral, nor the incentives as punishments/rewards. Therefore, this experiment tested whether we would still observe our effects in a social context that is divorced from moral concerns, while testing and controlling for the extent to which participants view the incentives as punishments or rewards.

Finally, this study provides further evidence that our results diverge from a pure loss aversion account by testing whether perceived norms and behavioral intentions remained different between conditions even when the value of the discount was twice that of the surcharge. Across several domains, a loss (e.g., a surcharge) has been estimated to

carry approximately twice the weight of an equally valued gain (e.g., a discount; Abdellaoui, Bleichrodt, & Paraschiv, 2005; Camerer, 2005). Thus, under a loss aversion account, offering a discount twice the value of a surcharge should be similarly motivating. However, even with this asymmetric structure, we expected that participants who were exposed to a surcharge would still perceive greater injunctive and descriptive norms, anticipate stronger norm-related emotional responses for failure to conform, and be more likely themselves to conform to the incentivized behavior.

Method

Eight hundred Mechanical Turk workers from the United States (43.5% female, 0.3% other, M_{age} = 33.8 years, SD_{age} = 10.6 years) participated the day before the American holiday of Thanksgiving. Participants were randomly assigned to condition (incentive framing: surcharge vs. discount) in a between-subjects design. They then viewed a flyer for a Thanksgiving Turkey Trot 5k Run, where dressing up like a turkey was optional but encouraged. Participants in the discount [surcharge] condition read, "Participants who do the 5k [not] dressed in a turkey costume will receive $4 off [pay an additional $2 on] their race ticket" (see Appendix for stimuli).

After reading the scenario, participants responded to a measure of injunctive norms assessing how expected it is to carry out the incentivized behavior: "How much do you think it's an expectation to dress up for the race?" (1 = Not at all to 7 = Very much). Next, we captured their anticipated emotional reactions. Whereas in Studies 1 and 2 we measured

anticipated embarrassment and guilt in multi-item scales that implicitly measured these focal emotions, this study assessed whether participants would explicitly report anticipating feeling guilt and embarrassment. Specifically, participants indicated how embarrassing it would be to not dress up for the race and how guilty they would feel if they did not dress up for the race (1 = Not at all to 7 = Very much). Next, they responded to an additional item designed to capture their perceptions of the normative behavior: whether "the norm (standard) for other participants" was to dress or not dress in a turkey costume.

Participants then imagined that several of their friends decided to run with them, and responded to three items assessing their intention to dress up. These items were adapted from previous research measuring behavioral intentions (Moon, Chadee, & Tikoo, 2008): "Given a choice, my friends will choose to dress up in a turkey costume for the run"; "There is a strong likelihood that I will dress up in a turkey costume for the run"; and, "I will recommend to my friends that we dress up in turkey costumes for the run" (1 = Do not agree at all to 7 = Agree completely). These items demonstrated high internal consistency ($\alpha = .89$) and were averaged together to create an index of behavioral intention.

In order to properly address alternative explanations, we followed the method used by Mulder (2008) by assessing perceived incentive size and using it as a covariate in all analyses. Discount [surcharge] participants answered, "How big does $4 off for dressing up [a $2 charge for not dressing up] feel?" (1 = Not at all to 7 = Very).

Next, we assessed whether participants viewed the discount and surcharge as a reward and punishment, respectively. We posit that surcharges (vs. discounts) will project

stronger social norms regardless of whether individuals view them as punishments or rewards. As such, participants allocated 100 points across a set of potential reasons why the race organizers would offer $4 off for dressing up [charge $2 for not dressing up], with the most points given to the reasons they think are most likely. The potential reasons were: "they want to reward participants for dressing up," "they want to punish participants for not dressing up," "they want to make more money," "they care about how much spirit people show," "they expect participants to want to show spirit," and "a different reason."

Next, participants completed the Attention-to-Social-Comparison-Information (ATSCI) scale (α = .88; Lennox & Wolfe, 1984), measuring one's tendency to conform to the behaviors of others. This scale includes items such as, "At parties I usually try to behave in a manner that makes me fit in" and "It's important to me to fit into the group I'm with." We predicted that this measure would moderate participants' responses to surcharges versus discounts, such that the incentive framing would influence only the individuals who care the most about fitting in. Finally, participants provided demographic information.

Results

Norm Perceptions. Supporting the idea that participants indeed perceived a stronger injunctive norm of dressing up in the surcharge condition, surcharge (vs. discount) participants reported that there was a significantly stronger expectation that runners would dress up (surcharge: M = 5.44, SD = 1.44 vs. discount: M = 4.55, SD = 1.59; b = .99,

t(797) = 9.22, p < .001, 95% $CI_{difference}$ = [.78, 1.20], d = .66). Participants were also significantly more likely to report that "the norm" was to dress up like a turkey for the race rather than not dress up (surcharge = 80.0% vs. discount = 64.7%; b = .94, z = 5.53, p < .001, 95% CI [9.24%, 21.46%], φ = .17).

Norm-Related Emotions. Surcharge participants anticipated feeling more embarrassed (surcharge: M = 3.20, SD = 1.74 vs. discount: M = 2.59, SD = 1.67; b = .73, t(797) = 6.12, p < .001, 95% $CI_{difference}$ = [.50, .97], d = .43) and guiltier than discount participants (surcharge: M = 3.18, SD = 1.95 vs. discount: M = 2.79, SD = 1.59; b = .54, t(797) = 2.90, p = .004, 95% $CI_{difference}$ = [.28, .80], d = .29) for not dressing up, consistent with the previous studies.

Intention. In line with our predictions, surcharge participants indicated significantly higher intention to dress as a turkey (M = 4.48, SD = 1.78) than did discount participants (M = 4.29, SD = 1.80; b = .41, t(797) = 3.43, p < .001, 95% $CI_{difference}$ = [.18, .65], d = .23). Thus, the surcharge again more strongly motivated intention to engage in the incentivized behavior.

Moderation by Conformity Tendency. We next examined intention to dress in a turkey costume as a function of condition (surcharge vs. discount), individual tendency to conform to others' behavior, and their interaction, in a linear regression with incentive size perceptions as a covariate.[14] We mean-centered conformity tendency and effect-coded condition (1 vs. -1 for surcharge vs. discount) for proper interpretation of the main effects

[14] Without the covariate, the main effect is directionally consistent but marginal (*b* = .10, *t*(796) = 1.57, *p* = .12) and the interaction remains significant (*b* = .20, *t*(796) = 2.45, *p* = .01), with the floodlight analysis revealing a similar pattern.

(Aiken & West, 1991). This analysis revealed a significant main effect of condition (b = .20, t(795) = 3.35, p < .001, 95% CI = [.08, .32], d = .22), whereby surcharge participants reported significantly higher intentions, and a significant main effect of conformity tendency (b = .38, t(795) = 5.04, p < .001, 95% CI = [.23, .53], d = .34), whereby individuals higher in conformity were more likely to dress up, qualified by a significant interaction (b = .18, t(795) = 2.47, p = .01, 95% CI = [.04, .33], d = .21).

To examine this interaction, we conducted a floodlight analysis (Spiller et al., 2013), applying the Johnson-Neyman procedure to identify the range(s) of conformity tendency (ranging in value from 1 = low to 6 = high) for which the simple effect of condition was significant. This analysis revealed that there was a significant positive effect of being in the surcharge condition on intention for any value of conformity tendency greater than 3.24 (71.25% of participants; b = .26, t = 1.96, p = .05), but no effect for any value of conformity tendency less than 3.24. Thus, participants who tend to conform to their peers' expectations were significantly more likely to dress up under the surcharge than under the discount, but participants who are less influenced by their peers' expectations were equally likely to dress up in both conditions. See Figure 3.4 for this pattern.

Perceived Purpose of Incentive. Finally, to test whether participants perceived different reasons for the surcharge versus discount incentives, we examined how they allocated "points" to each rationale.[15] In both conditions, participants allocated the greatest number of points on average to the rationale that "they want to reward participants for

[15] Note that 6% of participants did not allocate a total of 100 points and were not included in the primary analyses, although including them does not change the pattern nor the significance of the results.

dressing up" (surcharge: M = 25.71, SD = 22.46 and discount M = 33.56, SD = 22.52), suggesting that participants viewed both incentives primarily as rewards. Consistent with expectations, participants assigned more points to the reward rationale in the discount than the surcharge condition (b = 8.76, t(743) = 5.25, p < .001, 95% CI = [5.48, 12.03], d = .39). Further, while participants in both conditions assigned few points overall to a punishment motive, surcharge participants assigned more points to punishment (M = 6.59, SD = 12.91) than did discount participants (M = 2.72, SD = 7.44; b = 4.58, t(743) = 5.94, p < .001, 95% CI = [3.07, 6.10], d = .44). Importantly, the effect of condition on behavioral intention remains significant when controlling for both of these point allocations (p < .001 and neither allocation significantly predicted intention, ps > .26), suggesting these motives alone are not entirely responsible for participants' responses. This pattern further suggests that people may view surcharges more as punishments relative to discounts, but that overall, they are not considered a strong form of punishment, at least in this context. Point allocations for the other rationales are included in the Appendix.

Discussion

Study 3 shows that under a surcharge loss-equated at half the size of a discount, participants were still more likely to perceive stronger social norms and anticipate stronger emotional reactions to violating these norms, which led to higher intentions to carry out the incentivized behavior. Further, this effect was moderated by conformity tendencies: participants who tend to conform to their peers expressed higher intention to engage in a

behavior incentivized with a surcharge than with a discount, but participants low in the tendency to conform expressed equal intention across incentive frames. This finding further corroborates the role of norm leakage in the influence of incentives on behavior.

This study also demonstrates that surcharges and discounts may align with a punishment and reward frame, but that it is unlikely we need to be concerned with some of the downsides of punishments such as reactance (e.g., Balliet, Mulder, & Van Lange, 2011), at least in this context. Further, these findings suggest that surcharges (vs. discounts) signal social norms not only in contexts where the incentivized behavior may be construed as moral (e.g., green behaviors like reusable mug and bag usage), but also in contexts where the incentivized behavior has no moral flavor (e.g., dressing up). A pretest verified this assumption, finding that participants did not judge others who fail to conform to an incentive in this context as immoral (see Appendix for details). However, a clear limitation of this study is its hypothetical nature. Therefore, in Study 4, we test whether exposure to an incentive can influence real behavior, particularly once the incentive is removed.

STUDY 4: SURCHARGES CHANGE DOWNSTREAM BEHAVIORS

Study 4 investigated a new domain, health promotion, and assessed whether exposure to a surcharge incentivizing a behavior at one point in time could influence whether individuals would carry out that behavior in the same location, but when the incentive was no longer in place. Participants were given an opportunity to make a

purchase and were offered an incentive, framed either as a surcharge or a discount, for using hand sanitizer before purchasing. We tested if participants who initially witnessed a surcharge (vs. discount) were more likely to spontaneously carry out the encouraged behavior at a second point in time, when no incentive was present.

This design also allowed us to examine the downstream consequences of merely being exposed to these incentive policies, without necessarily being monetarily impacted by the incentives. All participants learned about the presence of a surcharge or a discount, but some may have chosen not to purchase the item associated with this incentive. Thus, these individuals heard about the policy, but were not financially affected by it. We tested whether these individuals still exhibited an increased tendency to carry out the initially incentivized behavior under a surcharge (vs. discount), consistent with our norms account.

Method

Two hundred ninety-four undergraduates (56.4% female; .7% other; 4.8% undisclosed; M_{age} = 21.2 years, SD_{age} = 2.2 years) from a large West Coast University participated in a laboratory study in return for course credit.[16] Participants were randomly assigned to condition (incentive framing: surcharge vs. discount) in a between-subjects design. Participants came into the lab, signed in, and were given $.50 as a thank-you for their participation in the study.

[16] One participant was excluded because she wished to wash her hands, but had a cut on her hand and thus could not use the hand sanitizer; a second participant was excluded because the research assistant did not inform this participant about the chance to take a cookie.

In the first phase of the study, participants proceeded one at a time to a room where a research assistant (RA) and a confederate waited, both of whom were not informed about the study's purpose nor the hypotheses. During each participant's interaction in this room, a confederate sat at a computer facing the participant and was instructed to appear to be doing homework, but also to periodically glance at the participant. The confederate's presence was intended to increase participants' perceptions that they were in a social setting with a member of their peer reference group (another college student). The RA, seated behind a table, asked the participants to view two school t-shirts hanging on the wall. Participants indicated which t-shirt they thought best represented the school community. This interaction was designed to subtly encourage participants to think about their social environment, while also masking the true focus of the study.

Next, the RA drew participants' attention to the table in front of them, which held several packs of gum and four signs (see Appendix for materials). One sign announced a sale on chewing gum, with a picture of the gum and its price (45 cents in the discount condition, and 40 cents in the surcharge condition).[17] A second sign informed participants of an incentive related to the gum, the text of which varied by condition. In the discount condition, this sign read, "You'll get a 5 cent discount if you use hand sanitizer! Help protect yourself and keep [school name] safe for everyone!" In the surcharge condition, this sign instead read, "You'll pay a 5 cent surcharge if you don't use hand sanitizer! Help protect yourself and keep [school name] safe for everyone!" A third and fourth sign on the

[17] The price of the gum was structured so that the eventual price would be the same in both conditions: it cost 40 cents for all participants who used the hand sanitizer (after a 5-cent discount; or without a 5-cent surcharge) and 45 cents for participants who did not (without a 5-cent discount; or after a 5-cent surcharge).

table were intended to present a rationale for the incentive. One sign included a picture of people holding hands and read, “We’re all in this together! Stop Germs from Spreading. Clean your Hands.” The other sign had a picture of a toilet and someone using hand sanitizer and read, “Public keyboards can contain more germs than a toilet seat! Protect yourself & others!” The hand sanitizer was located on a separate table in the room.

To initiate this brief interaction, the RA said,

> “We have Extra gum for sale at a special low price – 1 for 45 [40] cents! In addition, we are letting all lab participants know we’re making an effort to reduce the spread of germs across campus. As it turns out, public keyboards can be one of THE dirtiest surfaces and a key way in which germs are spread. So, to encourage participants to protect themselves and reduce the spread of germs across campus, the lab is offering a 5 cent discount [charging an additional 5 cents] on the price of the gum for participants who use the hand sanitizer [don’t use the hand sanitizer] before purchasing.”

The RA recorded whether or not the participant purchased the gum (yes/no), whether or not they used hand sanitizer (yes/no), and the number of pumps of hand sanitizer they used.[18]

Participants were then taken to the computer lab where they completed 30 minutes of filler survey material and unrelated studies (more details are in the Appendix), as in a typical session in the lab. Embedded in a filler survey, participants answered on a 10-point

[18] Of participants who used the hand sanitizer, nearly all used exactly 1 pump (89.4%). The number of pumps participants used was equal across conditions ($\chi^2(1, N = 179) = 2.06$, $p = .36$).

scale, “Out of every 10 participants, how many do you think used the hand sanitizer before coming to the computer lab?” and provided demographic information.[19]

After this half hour of unrelated studies, participants returned individually to the same room used during phase 1 of the study. The RA handed the participants a sheet of paper and asked them to complete a “memory test” as the final part of the study. The paper had a picture of the two t-shirts on it from phase 1 and asked participants to circle the shirt they had voted for. This task was designed to justify why participants needed to return to the original room. When finished, participants handed their sheet to the RA and signed out of the lab session. On the table near the exit (across the room from the RA) was a tray of cookies along with a bottle of hand sanitizer. The RA dismissed each participant, telling them, “Thank you very much; you are free to go! There were extra cookies available from an earlier event here, so you can take one on the way out!” Thus, during the entirety of this second interaction, there was no mention of the hand sanitizer. The RA surreptitiously recorded whether the participant used the hand sanitizer (yes/no) as they left the lab.[20]

Results

[19] Participants also responded to the ATSCI scale. This measure did not moderate behavior in this experiment; we suspect this is the case because this individual difference scale was presented completely divorced of any social context. Further, the scale leans toward descriptive norm items, for which we found no effect in this experiment. This may serve as further evidence of the important role of injunctive norms in driving behavior in this setting.

[20] The RA also recorded how many pumps of hand sanitizer participants used. As in phase 1, of participants who used the hand sanitizer, nearly all used exactly one pump (88.6%) and the number of pumps used was equal across conditions ($\chi^2(1, N = 44) = .72$, $p = .70$). The RA also recorded whether the participant took a cookie and whether the participant took napkins with their cookies. Participants were equally likely across conditions to take a cookie (discount: 71.4% vs. surcharge: 65.3%, $\chi^2(1, N = 294) = 1.27$, $p = .26$), and to take a napkin (discount: 53.7% vs. surcharge: 51.7%, $\chi^2(1, N = 294) = .12$, $p = .73$). Thus, exposure to the incentive in phase 1 influenced only the specific behavior that was incentivized, and not other behaviors.

Phase 1. Participants purchased gum at equal rates in both conditions (surcharge: 58.5% vs. discount: 54.4%, $\chi^2(1, N = 294) = .498$, $p = .48$, 95% CI [-15.31%, 7.25%], $\varphi = .04$). It turns out that the calibration of the incentives led to a ceiling effect on phase 1 hand sanitizer usage among participants who purchased gum (surcharge: 94.2% vs. discount: 96.3%; $\chi^2(1, N = 166) = .38$, $p = .54$, 95% CI [-6.00%, 10.20%], $\varphi = .05$) and equal usage among those who did not (discount: 16.4% vs. surcharge: 16.4%, $\chi^2(1, N = 126) < .001$, $p > .99$, 95% CI [-13.02%, 13.32%], $\varphi < .001$).

Phase 2. The focus of this experiment was on phase 2 hand sanitizer usage. As predicted, even with the incentive no longer in place, surcharge participants were significantly more likely to spontaneously use the hand sanitizer in phase 2 (19.7%) than were discount participants (10.2%; $\chi^2(1, N = 294) = 5.24$, $p = .02$, 95% CI [1.39%, 17.84%], $\varphi = .13$). Thus, encountering an incentive framed as a surcharge at one point in time nearly doubled the proportion of participants who carried out the encouraged behavior later on, when it was no longer incentivized—and not even mentioned.

Importantly, the effect of condition on hand sanitizer usage in phase 2 was driven by the participants who had not purchased any gum in phase 1 (43.5% of participants). These individuals were merely exposed to either a surcharge or discount, and were not financially affected by it. Yet, even among these participants, those who had observed a surcharge in phase 1 were significantly more likely to spontaneously use the hand sanitizer in phase 2 (36.1%) than were those who had observed a discount (16.4%; $\chi^2(1, N = 128) = 6.44$, $p = .01$, 95% CI [4.48%, 34.48%], $\varphi = .22$). This pattern suggests that merely being

exposed to a surcharge, even while not being financially affected by it, can leak social norm information and significantly alter behavior thereafter.

Survey Measure: Descriptive Norm. Participants perceived similar rates of hand sanitizer usage by their peers in both conditions (surcharge: M = 5.80, SD = 3.17 vs. discount: M = 5.74, SD = 3.31; t(278) = .17, p = .87, 95% $CI_{difference}$ = [-.83, .70], d = .02), contrary to our expectations. However, participants' estimates of others' behavior was reflective of their own actual usage: participants who used hand sanitizer reported thinking that significantly more of their peers also did so (b = 4.05, t(278) = 12.87, p < .001, 95% $CI_{difference}$ = [3.43, 4.67], d = 1.55).

Discussion

Study 4 provides behavioral evidence that exposure to a surcharge (vs. discount) can leave a lingering impact: the use of surcharges (vs. discounts) to encourage hand sanitizing doubled the proportion of individuals enacting this behavior at a later point in time, when no incentive was in place and the behavior was not even mentioned. Further, this effect arose even among individuals who were not financially impacted by the original incentives, providing evidence that mere exposure to an incentive frame can leave lasting impressions even on individuals who do not experience the surcharges or discounts themselves.

Despite the social nature of our effect, we did not find that surcharge participants perceived a significantly stronger descriptive norm of hand sanitizer use. This result may

be explained by false consensus, the phenomenon in which people's egocentric tendencies lead them to expect similarities between themselves and others and overestimate the commonality of their own behaviors (e.g., Mullen et al., 1985; Ross, Green, & House, 1977). Indeed, participants may very well have inferred that others had simply behaved just as they had (Tankard & Paluck, 2016). As this question was asked directly after their own usage (or non-usage) of the hand sanitizer, this item may not be a clean measure of a perceived descriptive norm, as it is possible that they were merely projecting their own behavior onto their peers.

GENERAL DISCUSSION

Shifting social norms is a valuable organizational tool, as norms can influence talent and performance (Swaab & Galinsky, 2015) and even organizational productivity, spontaneity, and success (George & Jones, 1997, Stewart et al., 2012; De Jong, Bijlsma-Frankema, & Cardinal, 2014). Across four studies, we demonstrate that framing an incentive as a surcharge, as compared to a discount, leaks stronger injunctive and descriptive social norms, signaling both greater social expectations to perform and higher prevalence of the incentivized behavior. The prospect of violating these perceived social norms if one fails to perform the incentivized behavior elicits higher anticipated emotional reactions of embarrassment and guilt which, in turn, increase one's likelihood of carrying out the perceived normative behavior.

Importantly, this framework implies a set of novel predictions as to the function and consequences of incentives. First, we show that incentives framed as surcharges have a lingering effect, influencing behaviors even in locations and at times when no incentive is in place (Studies 2 and 4), an effect mediated by perceived norms and norm-related emotions (Studies 1 and 2). Second, because their framing leaks information about social norms, surcharges can cause people to anticipate feeling embarrassed and guilty if they fail to perform the incentivized behavior (Studies 1-3). Third, by invoking norms, incentives framed as surcharges (vs. discounts) can influence downstream behaviors of both individuals who were previously financially affected by the incentive as well as those who merely observed the incentive (without being financially affected by it; Study 4). Finally, we find that surcharge (vs. discount) incentives impact people differently depending upon their individual sensitivity to social norms, with surcharges most strongly affecting the individuals who care the most about conforming to their peers (Study 3).

We show the robustness of our findings by demonstrating that they hold across a range of domains and contexts, including both moral (e.g., "green" behaviors like bringing reusable grocery bags) and non-moral contexts (e.g., dressing up for a themed event), as well as for both hypothetical and real behaviors. Further, our results persist both when there is and when there is not a plausible cost-passing justification for the incentive. That is, people may perceive that retailers implement surcharges when they want to pass along the cost of providing a material product (e.g., the cost for them to provide plastic grocery bags), or the cost of providing additional services (e.g., additional minutes of childcare; Gneezy & Rustichini, 2000). However, we find that surcharges leak stronger social norms

and have similar consequences for behavior in contexts both with (e.g., when supplying grocery bags or coffee cups) and without (e.g., when encouraging individuals to dress in costume or use hand sanitizer) the potential for cost-passing, increasing the application of our findings.

Theoretical Contributions

The present work makes notable theoretical contributions to research on social norms and organizational behavior, incentives, and policy. First, we contribute to the organizational behavior and social norms literatures by introducing a novel tool that can be used to harness the well-established power of norms to effectively change behaviors across many domains (e.g., Berkowitz, 2004; Cialdini & Trost, 1998; Cialdini, Reno, & Kallgren, 1990; Larrick, Soll, & Keeney, 2015). Prior research suggests that individuals infer social norms via several channels, including the behavior of other individuals, the behavior of groups, and signals from institutions, such as organizational rules or public policies (Tankard & Paluck, 2016). In this work, we demonstrate that the framing of incentives is a novel channel through which injunctive and descriptive social norms are leaked, signaling the approved behaviors in a community, as well as how others tend to behave. More broadly, our work opens the door to investigate whether other features of organizations and the marketplace have the potential to similarly signal social norm information, and thus potentially serve as tools for changing behaviors.

Second, the current research also adds to the policy and morality literatures by building on previous work suggesting that the structure of a law or policy can lead to inferences about the morality of a behavior. Specifically, using social dilemma games, Mulder (2008) finds that participants judge violators as more immoral in a setting where punishments are given for non-compliant behaviors than when rewards are given for compliant behaviors, which she suggests occurs because punishments signal that a behavior is morally obligatory. While injunctive norms may convey a moral principle, we suggest that this diverges from whether a behavior is inherently moral (e.g., donating to charity), from moral concerns (e.g., if I do not perform this behavior, others will think I am a bad person), and from moral judgments (e.g., that someone who does not behave in this manner is morally unfit). Thus, the current research expands this previous work on the roles of policy framing and morality by demonstrating that surcharges leak general injunctive and descriptive social norms, not just moral norms, across a broader context of consumer and organizational behaviors. Further, we demonstrate that while surcharges and discounts map on to punishments and rewards and indeed may be perceived as such in some contexts, this is certainly not always the case and is not a necessary component of our theory. This suggests that policy framing may not be limited to a punishment/reward domain and may have larger implications than has been investigated in prior research (Evers et al., 2016; Mulder, 2008, 2016).

Finally, this work contributes to the incentives literature by demonstrating that the framing of an incentive has greater implications than previously believed, and by illustrating the potential behavioral ripple effects of incentives. Prior research investigating

surcharges has attributed their superior effectiveness over discounts mainly to loss aversion (Homonoff, 2015; Nasiry & Popescu, 2011; Poortinga, 2017), under the assumption that surcharges are perceived as losses and discounts as gains. However, loss aversion draws on the monetary valuation of the incentive itself. Accordingly, loss aversion can help describe responses to active incentives in which an individual is directly impacted by the feeling of a monetary loss. However, our research demonstrates several important consequences that would not be predicted by this monetary valuation asymmetry, and instead would only arise if incentive framing does indeed leak social norms. First, we demonstrate that failing to carry out the incentivized behavior led individuals to anticipate emotions associated with the violation of a social norm (e.g., embarrassment and guilt). Second, mere exposure to a surcharge at one point in time led to downstream changes in behavior where the incentive was not in place—even among people who merely observed the incentive and were not monetarily impacted by it, an outcome consistent with research demonstrating that shifts in perceived social norms can lead to lasting behavioral effects (Neighbors, Larimer, & Lewis, 2004; Schultz et al., 2007). Third, even with loss-equated incentives, individuals responded more strongly to a surcharge because of the social norm information it carried. Finally, individuals who exhibited a natural propensity to be influenced by peer expectations responded more strongly to surcharges than discounts, while individuals who cared less about conforming to peer expectations responded similarly to both—a result that would not be predicted by individuals merely responding to the financial value of the incentives. In summary, our findings offer novel evidence of an important undiscovered consequence of incentives.

Managerial Implications

Incentives are a popular organizational tool used to encourage engagement in a wide variety of behaviors, ranging from environmentally friendly actions (e.g., using renewable energy, purchasing a hybrid car, and recycling) to health-related behaviors (e.g., wellness programs that offer incentives for smoking cessation, weight loss, and management of chronic diseases; Cigna, 2017). Thus, in addition to the theoretical contributions of this research, our work carries key practical implications relevant to organizational behavior.

First, our findings suggest that the manner in which organizations and policymakers frame these incentives may shift perceptions of social norms, leading to carryover behaviors and potentially habit formation (Ouellete & Wood, 1998)—even once the incentive ends. Specifically, we demonstrate that incentive policies framed as surcharges have the potential to influence behaviors downstream both in the same location as well as carry over to other locations where similar social norms are expected.

Second, framing an incentive as a surcharge can give managers a "bigger bang for their buck." As shown in Study 3, a surcharge can lead to greater compliance than a discount even twice its price. Thus, managers wishing to best incentivize positive workplace behaviors may consider that the way they frame a policy may have a larger impact than previously realized. For example, if they choose to frame an incentive as a surcharge, in the right context, they may successfully signal that a behavior is more

normative. This could better motivate organizational citizenship behaviors such as meeting attendance, corporate social responsibility activities such as volunteering, or uptake of flexible work hours. However, it is important to consider whether employees will perceive such surcharges as aligned with the community and organizational beliefs or as unjust, a possibility that could elicit reactance. A further consideration is that surcharges (if perceived as "stick" policies) may increase perceptions that a person or behavior is viewed negatively by the organization or policymaker, potentially stigmatizing certain groups and increasing distaste for the policy (Tannenbaum et al., 2013). The current studies examining small-scale surcharge and discount incentives (e.g., 10 cents) did not find evidence for these concerns, but they should be considered.

Finally, our findings suggest that signaling a social norm through the framing of an incentive may influence a larger audience than would be expected by a financial-motivation account, as individuals who are not motivated by a 5-cent, 10-cent, or even two-dollar incentive may still care about conforming to what they perceive to be a social norm. Further, the framing of the incentive may spill over to a larger audience by influencing individuals who observe the incentive but are not directly financially impacted by it. Indeed, we demonstrate that merely encountering a surcharge (and not directly experiencing it) led to greater incidence of the previously incentivized behavior. In sum, our findings suggest a new lever for influencing norm perceptions, with the potential to influence a wider audience and cause behaviors to linger and carry over to locations where no incentive is in place.

Limitations and Directions for Future Research

While this investigation outlined robust support for our account, naturally the power of surcharges is not without limits. Our theorizing rests on people inferring that the surcharge is acceptable and put in place by a reasonable policymaker. Thus, an important boundary condition for our work may be that the incentive has to seem legitimate and justified. Beginning to impose an extra charge that goes against any established, well-known group norms or that seems unfair may provoke reactance, and would likely not have the intended effect on perceived norms or behavior (Payne, 2001; Shah et al., 2014; Thaler, 1980). Thus, future research could investigate whether the perceived appropriateness of the incentive moderates the impact of incentive framing on perceived social norms and behavioral responses.

An additional potential boundary condition concerns the magnitude of the incentives. In the present research, we investigated relatively small pricing incentives (a few cents to a few dollars). It is possible that at larger incentive values, the motivational superiority of surcharges over discounts may cease to exist, or may even reverse, or that financial considerations might dwarf the social norms mechanism we identified. Further research could explore such possibilities. Future research could also examine whether incentivizing a behavior affects other different, but related, behaviors. For example, if framing an incentive as a surcharge encourages an individual to bring her own bags to the grocery store, it may serve as a self-signal (Bem, 1972) that she cares about the environment, leading to potential shifts in her environmental attitudes and other

environmentally friendly behaviors (Poortinga, Whitmarsh, & Suffolk, 2013; Thomas, Poortinga, & Sautkina, 2016). On the other hand, increases in environmentally responsible behaviors such as using fewer plastic bags may give individuals license to misbehave in other ways (Khan & Dhar, 2006; Karmarkar & Bollinger, 2015), pointing to an interesting potential avenue for research.

Furthermore, the current research investigated potential spillover effects of exposure to an incentive at one point in time. Specifically, we find that exposure to a surcharge (vs. discount) can lead individuals to carry out the encouraged behavior at a later point in time when no incentives are in place. However, the time scale of our studies was relatively short (a 30-minute gap in consequential Study 4 and a hypothetical 1-week gap in scenario Study 2). Future research may wish to examine longer-term interventions and test potential longer-lasting impacts of initial exposure to surcharge incentives. Finally, we examine differences in the psychology of loss and gain frames and how they invoke norm inferences, leading to differential effects on emotions and behavior. It is possible that this psychology could contribute to the general phenomenon of loss aversion; however, specific work would be needed to test this conjecture, providing a potential avenue for future research.

In sum, our studies add to the growing body of knowledge about the nature of incentives and the inferences people draw from them by demonstrating that merely framing an incentive as a surcharge, rather than a discount, can influence perceived social norms and lead to meaningful changes in behavior. With this knowledge, careful consideration should be given to the behavioral nudges we use in management, marketing, policy-

making, economics, psychology, and beyond, as the mere framing of incentives is more powerful than previously realized.

ACKNOWLEDGEMENTS

We are grateful to the editors and reviewers of this paper for their helpful input.

Chapter 3, in full, is a reprint of previously published material as it appears in Organizational Behavior and Human Decision Processes, 151: 118—131, Lieberman, Alicea, Kristen Elizabeth Duke, and On Amir. The dissertation author was one of the primary investigators and authors of this paper.

FIGURES

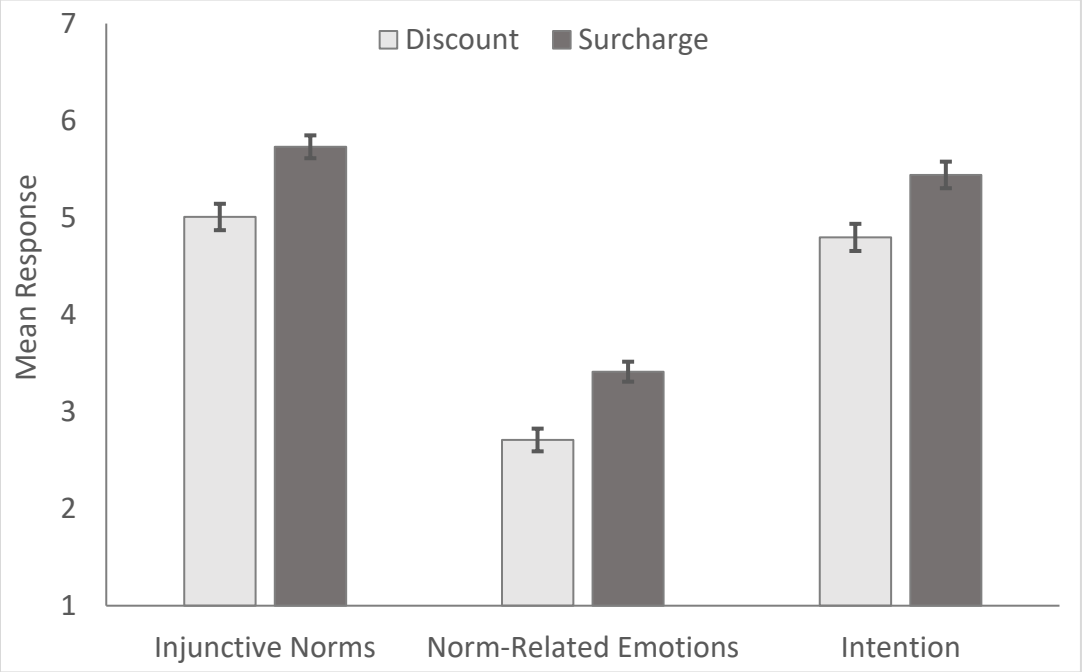


Figure 3.1. The effects of incentive framing on perceived injunctive norms, embarrassment and guilt (averaged), and intention to bring a reusable bag, Study 1. Error bars represent ± 1 SE around the mean.

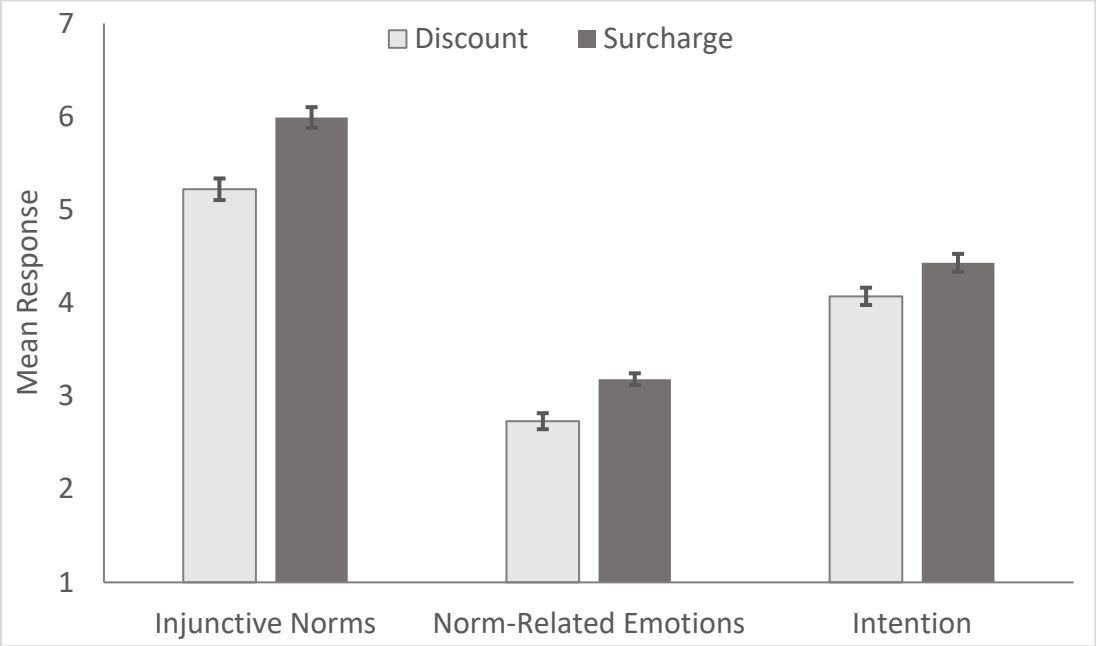


Figure 3.2. The effects of incentive framing on perceived injunctive norms, embarrassment and guilt (averaged), and intention to bring a reusable bag, Study 2. Error bars represent ± 1 SE around the mean.

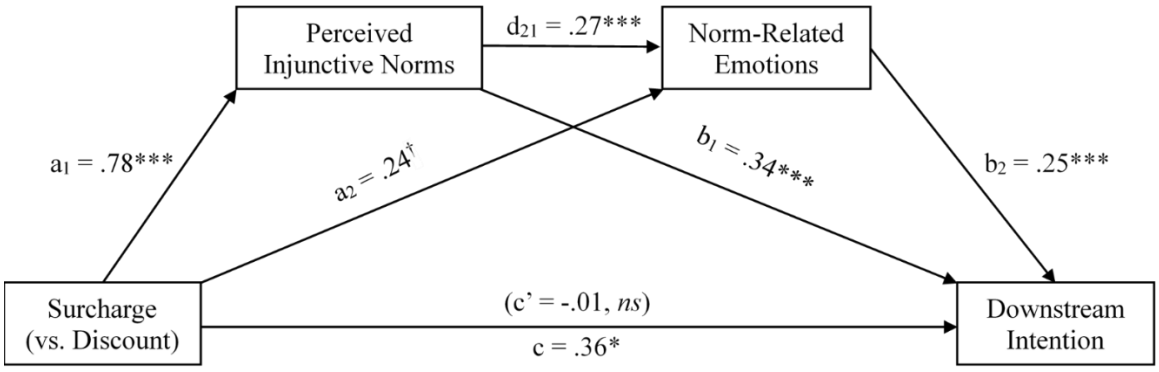


Figure 3.3. The effects of incentive framing on downstream intention to bring reusable bags, through perceived injunctive norms and norm-related emotions, in Study 2. The path coefficients are unstandardized betas. The value in parentheses indicates the direct effect of surcharges on downstream intention after accounting for the two mediators. Norm-related emotions are an average of guilt and embarrassment. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$

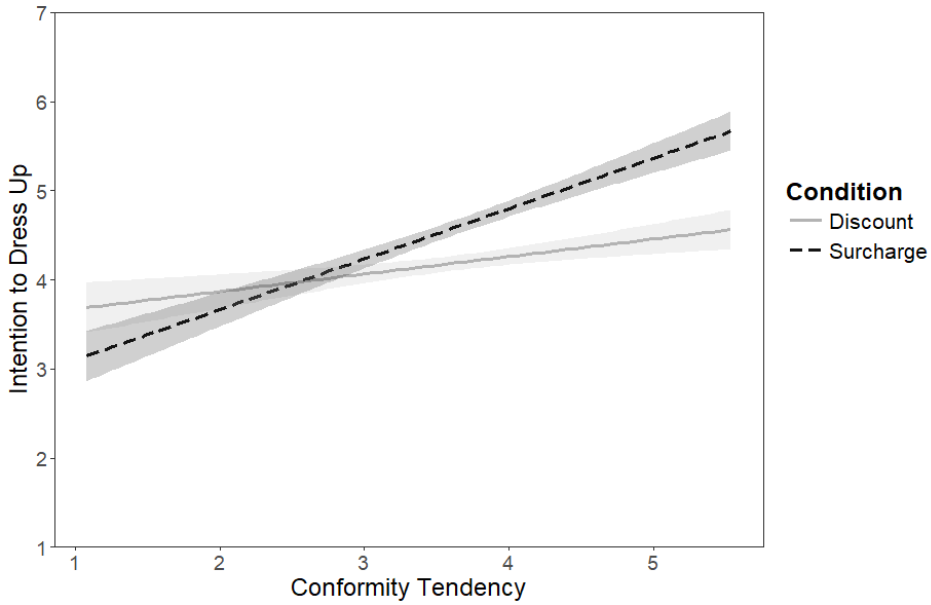


Figure 3.4. Regression-predicted intention to dress up by incentive condition, moderated by conformity tendency, in Study 3. Incentive size perceptions was included as a covariate: regression-predicted values were computed at the mean level of incentive size perceptions. Shaded areas reflect ± 1 SE around the predictions.

APPENDIX

Additional Discussion of Self-Conscious Emotions

In the manuscript, we provide a discussion of guilt and embarrassment, two self-conscious emotions that are action-oriented and have been shown to follow violations of social norms (Lewis, 1971; Tracy & Robins, 2004). However, there are two additional self-conscious emotions, shame and pride, that are less relevant to the contexts in this research, and thus were not included in our experiments. Nevertheless, here we provide a more thorough discussion of these two self-conscious emotions.

Shame

Shame involves a global evaluation of the self as violating "society's definition of what it means to be a worthy, competent, good, and respectable person" (p. 205-206, Edelstein & Shaver, 2007), and accordingly motivates extreme withdrawal and social avoidance (e.g., Tangney, Miller, Flicker, & Barlow, 1996). Shame is typically viewed as an intense emotional response to a behavior that violates deeper moral rules (Turiel, 1983), whereas our theory proposes that social norm signaling in these contexts does not necessarily involve moral concern.

Moreover, shame involves an appraisal about an individual's overall character (Lewis, 1971; Tracy & Robins, 2004), motivating avoidance behavior (Tangney et al., 1996). As such, imagining "undoing" the past to avoid shame elicits thoughts of changing aspects of oneself (Niedenthal, Tangney, & Gavanski, 1994), versus undoing aspects of one's behavior (as with guilt). For these reasons, shame is less relevant to our contexts, as we do not expect small incentives to fundamentally shift whether people perceive a dispositional flaw in their character (Hart & Matsuba, 2007) which would lead to behaviors of avoidance; rather, we expect that they project social conventions that drive individuals to seek to fit in.

Pride

Pride is a positive self-conscious emotion. Prior research contrasts two forms of pride that parallel the distinction between shame and guilt (e.g., Tracy & Robins, 2007). Like shame, hubristic pride arises following attributions about positive stable, global aspects of the self (e.g., Lewis, 2000), while like guilt, authentic pride involves a focus on specific behaviors (Tracy & Robins, 2004). And, while authentic pride is at times a motivating emotional force, it tends to involve an appraisal of the self as outperforming others, rather than matching others' behavior (as would be expected when one attempts to match the norm; Exline & Lobel, 1999; van Osch, Zeelenberg, & Breugelmans, 2018). Thus, we do not expect that the social norm signaled by an incentive frame will elicit strong experiences of pride in our contexts.

Study Stimuli, Study 1

In Study 1, participants were presented with visual stimuli of coffee shop signs differing by condition. Discount participants saw a sign that read, “Customers will get a 10 cent discount for bringing their own mugs!” while surcharge participants saw a sign that read, “Customers will be charged 10 cents for not bringing their own mugs!” (see Figure A3.1).

Discount

Surcharge

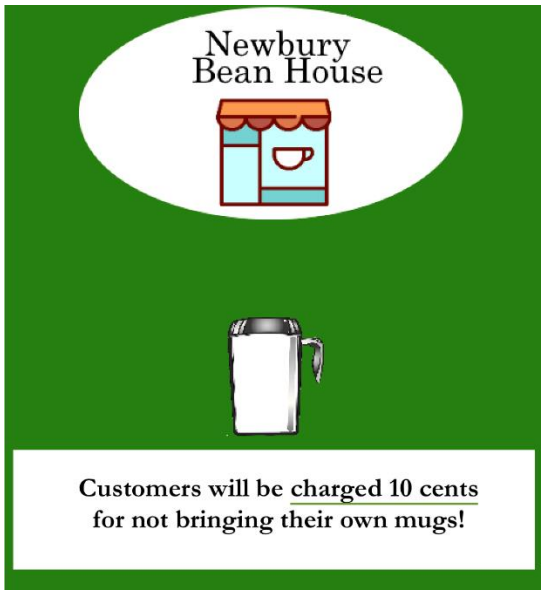


Figure A3.1. Flyers in each condition.

Serial Mediation Models, Study 1

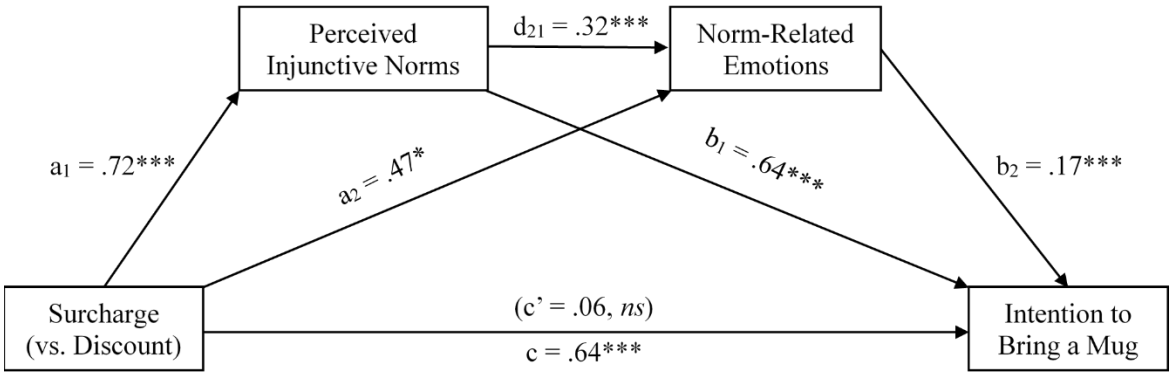


Figure A3.2. Injunctive norms as the first mediator. The effects of incentive framing on behavioral intention to bring a reusable mug, through perceived injunctive norms and anticipated norm-related emotions, in Study 1. The path coefficients are unstandardized betas. The value in parentheses indicates the direct effect of surcharges on intention after accounting for the two mediators. Norm-related emotions are an average of guilt and embarrassment. *** p < .001; ** p < .01; * p < .05

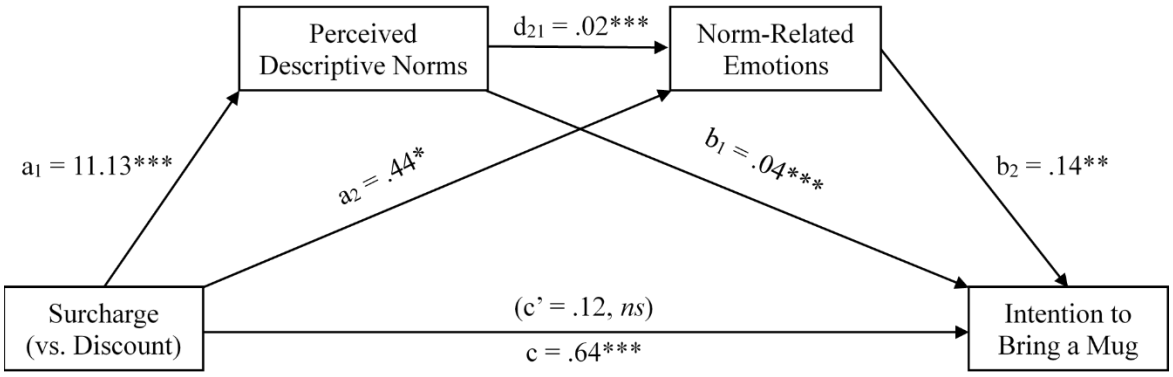


Figure A3.3. Descriptive norms as the first mediator. The effects of incentive framing on behavioral intention to bring a reusable mug, through perceived descriptive norms and anticipated norm-related emotions, in Study 1. The path coefficients are unstandardized betas. The value in parentheses indicates the direct effect of surcharges on intention after accounting for the two mediators. Norm-related emotions are an average of guilt and embarrassment. *** p < .001; ** p < .01; * p < .05. Serial paths $a_1 \times d_{21} \times b_2 = .04$, SE = .02, 95% CI: [.01, .07].

Replication of Study 1

In Study 1, a surcharge was imposed on participants who did not perform a behavior, while a discount was offered to participants who did perform a behavior. Thus, the surcharge condition involved inaction and the discount involved action. In this study, we replicate the effects of Study 1 while eliminating this difference. Specifically, the incentives here are designed such that the focal behavior involves an action in both conditions (i.e., "Customers will get a 10 cent discount for bringing their own bags" and "Customers will be charged 10 cents for using our bags"). We demonstrate that when both conditions are framed as requiring an action, surcharges again lead to stronger injunctive and descriptive norms and greater norm-related emotions surrounding the encouraged behavior.

Method

Two hundred Mechanical Turk workers (37.5% female, 0.5% other, M_{age} = 33.7 years, SD_{age} = 9.78 years) participated.[21] Participants were randomly assigned to one of two conditions (incentive framing: surcharge vs. discount) in a between-subjects design. All participants read, "Imagine that you have just moved to a new town, Newbury, and do not know anyone yet. You are looking for opportunities to get to know your new neighborhood. You remember that you need to pick up some groceries, so you head to the closest grocery store." On the next page, participants read that upon entry to the store, they saw a sign. They were presented with an image of this sign. Discount participants saw, "Bring your own reusable bags to reduce waste! Customers will get a 10 cent discount for bringing their own bags!" Surcharge participants instead saw, "Bring your own reusable bags to reduce waste! Customers will be charged 10 cents for using our bags!"

Following this description, participants answered two questions in counterbalanced order designed to measure both an injunctive norm (item 1) and a descriptive norm (item 2): "How much do you think the checkout cashier expects customers to bring their own bags?" (1= Not at all to 7 = Very much) and "Out of every 100 customers, how many do you think bring their own bags?" On the following page, participants responded to two questions in counterbalanced order measuring their anticipated emotional reactions: "How embarrassed would you feel if you did not bring your own bags to this store?" and "How guilty would you feel if you did not bring your own bags to this store?" (1 = Not at all to 7 = Very). Finally, participants provided demographic information.

Results

Injunctive Norms. Surcharge (vs. discount) participants believed that the checkout cashier had significantly higher expectations that customers would bring their own bags

[21] Twelve participants failed an attention check ("What was the incentive offered at the grocery store?" [10 cent discount, 5 cent discount, 10 cent charge, 5 cent charge, or I don't remember/I'm not sure]), indicating they were unaware of the incentive framing for their condition. Therefore, we excluded these participants from analyses, although including them does not change the pattern of results.

(surcharge: $M = 4.98$, $SD = 1.71$ vs. discount: $M = 3.99$, $SD = 1.64$; $t(186) = 4.04$, $p < .001$), indicating a stronger injunctive norm.

Descriptive Norms. Participants expected that significantly more customers would bring reusable bags in the surcharge condition ($M = 55.24$, $SD = 25.82$) than in the discount condition ($M = 40.55$, $SD = 25.21$; $t(186) = 3.95$, $p < .001$), indicating a stronger descriptive norm.

Norm-Related Emotions. Participants who read about a surcharge (vs. discount) anticipated feeling guiltier (surcharge: $M = 3.80$, $SD = 1.95$ vs. discount: $M = 3.19$, $SD = 1.93$; $t(186) = 2.14$, $p = .03$) and more embarrassed (surcharge: $M = 3.31$, $SD = 1.86$ vs. discount: $M = 2.67$, $SD = 1.78$; $t(186) = 2.40$, $p = .02$) for failing to bring a bag to the store.

Discussion

Even in a context where both incentives involved action, the surcharge still projected stronger injunctive and descriptive norms than did the discount, leading to consistent shifts in anticipated embarrassment and guilt.

Study Stimuli, Study 2

In Study 2, participants were presented with visual stimuli of neighborhood grocery store signs differing by condition. Discount participants saw a sign that read, “Customers will get a 10 cent discount for bringing their own bags!” while surcharge participants saw a sign that read, “Customers will be charged 10 cents for not bringing their own bags!” (see Figure A3.4).

Discount Surcharge

Figure A3.4 Flyers in each condition.

Serial Mediation Model, Study 2

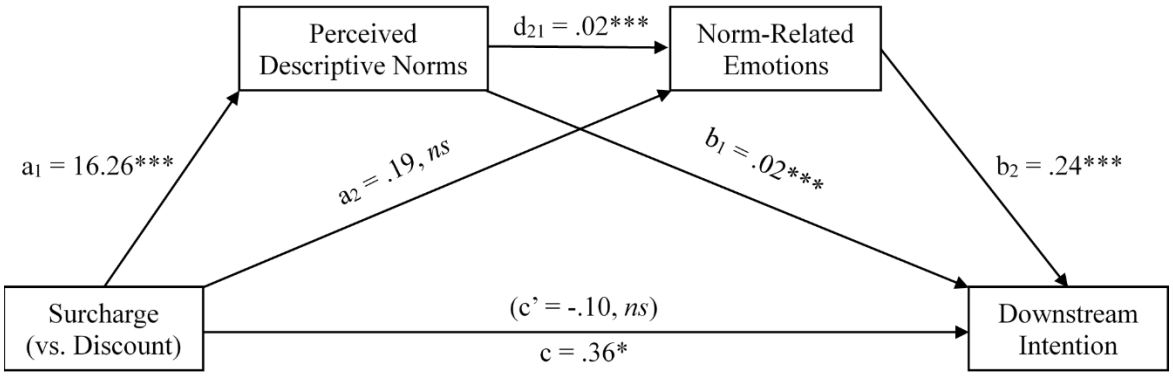


Figure A3.5. Descriptive norms as first mediator. The effects of surcharges on downstream intention to bring reusable bags, through perceived descriptive norms and norm-related emotions, in Study 2. The path coefficients are unstandardized betas. The value in parentheses indicates the direct effect of surcharges on downstream intention after accounting for the two mediators. Norm-related emotions are an average of guilt and embarrassment. *** $p < .001$; ** $p < .01$; * $p < .05$. Serial paths $a_1 \times d_{21} \times b_2 = .06$, SE = .02, 95% CI: [.03, .10].

Study Stimuli, Study 3

In study 3, in addition to reading text explaining the incentive, participants saw an image of the turkey trot flyer with text differing by condition. Participants in the discount saw a sign with text that read: “Come in costume! ‘Human’ participants pay an additional $2 on race tickets!” while surcharge participants saw a sign with text that read, “Come in costume! ‘Turkey’ participants get $4 off of race tickets!" (see Figure A3.6). In addition, right below the turkey trot flyer participants in both conditions saw an image of people dressed up in a turkey trot race (see Figure A3.7).

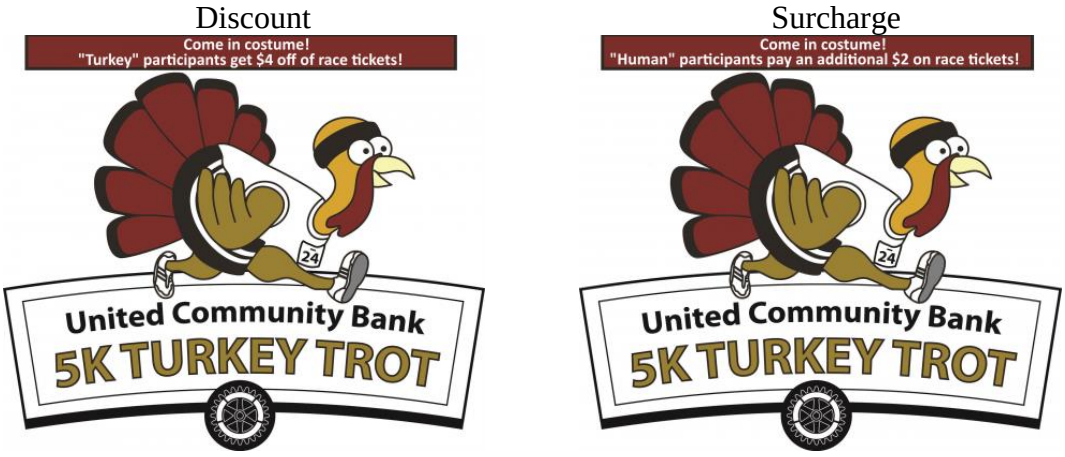


Figure A3.6. Signs announcing the incentive in each condition.

Figure A3.7. Image seen by participants in both conditions.

Pretest, Study 3

Past research suggests that sanctions framed as punishments (vs. rewards) signal a moral obligation, leading people to judge a violator of these perceived obligations as more immoral (Mulder, 2008). Our account of the effect of incentive framing is agnostic to any moral concerns, as discussed in the manuscript. As such, the scenario in Study 3 was designed such that the focal behavior had no moral flavor: dressing up in a turkey costume for a Thanksgiving-themed run.

To verify the assumption that this context was not morally linked, we conducted a pretest with 201 Mechanical Turk participants. Participants were randomly assigned to read either the discount or surcharge scenario and view the flyers from the main study (described in the main text, and shown in Figure 6). Following the scenario, participants read, "Imagine it's the day of the run, and you're lining up for it to begin. You're standing behind someone who is not dressed in a turkey costume. To what extent is this person acting immorally?" (1 = Not at all to 7 = Very much).

As expected, participants did not judge this violation to be immoral—77% of participants responded "not at all"—and surcharge and discount participants responded equally (surcharge: $M = 1.58$, $SD = 1.18$ vs. discount: $M = 1.48$, $SD = 1.24$; $t(199) = .56$, $p = .58$). This suggests both that this context is not viewed as morally linked, and that the incentive frame does not leak information about the moral obligation of the incentivized behavior itself.

Additional Point Allocations, Study 3

The main manuscript presents detailed analyses addressing two of the categories to which participants could allocate points (the reward and punishment motives). In Table A3.1, we present the allocations to all of the motives. Note that the effect of incentive framing on intention holds when controlling for all of these point allocations, as well.

Table A.3.1. Additional point allocations, Study 3. Values in the discount and surcharge columns represent means, with standard deviations in parentheses. Statistics present t-tests comparing the two rationales, although all differences hold when also controlling for the perceived size of the incentive as in the main manuscript.

Rationale	**Discount**	**Surcharge**	**Difference**
"They want to reward participants for dressing up"	33.56 (22.52)	25.71 (22.46)	$t(744) = 4.76$, $p < .001$
"They want to punish participants for not dressing up"	2.72 (7.44)	6.59 (12.91)	$t(744) = 5.03$, $p < .001$
"They want to make more money"	7.27 (15.37)	21.07 (24.68)	$t(744) = 9.19$, $p < .001$
"They care about how much spirit people show"	28.01 (19.77)	22.27 (18.99)	$t(744) = 4.05$, $p < .001$
"They expect participants to want to show spirit"	26.55 (17.82)	23.38 (19.89)	$t(744) = 2.29$, $p = .02$
"A different reason"	1.89 (8.95)	0.99 (5.29)	$t(744) = 1.67$, $p = .10$

Study Stimuli, Study 4

In Study 4, participants encountered four signs upon entering the room. One sign announced a sale on chewing gum, with a picture of the gum and its price (45 cents in the discount condition, and 40 cents in the surcharge condition; see Figure A3.8).

Figure A3.8. Sign announcing the sale of gum. The base price was displayed as 40 cents in the surcharge condition.

A second sign informed participants of an incentive related to the gum, the text of which varied by condition. In the discount condition, this sign read, “You’ll get a 5 cent discount if you use hand sanitizer! Help protect yourself and keep [school name] safe for everyone!” In the surcharge condition, this sign instead read, “You’ll pay a 5 cent surcharge if you don’t use hand sanitizer! Help protect yourself and keep [school name] safe for everyone!” (see Figure A3.9).

Discount Surcharge

Figure A3.9. Signs announcing the incentives.

A third and fourth sign on the table were intended to present a rationale for the incentive. One sign included a picture of people holding hands and read, "We're all in this together! Stop Germs from Spreading. Clean your Hands." The other sign had a picture of a toilet and someone using hand sanitizer and read, "Public keyboards can contain more germs than a toilet seat! Protect yourself & others!" (see Figure A3.10).

Figure A3.10. Signs providing rationale for the incentive.

Filler Survey Items, Study 4

Participants completed 30 minutes of filler survey material and participated in an unrelated study between phase 1 and phase 2. The filler survey items included: “How often do you wash your hands before eating?” (1 = Absolutely never to 6 = Absolutely every time); “To what extent are you currently dieting?” (1 = Not at all to 9 = Very much); “Do you have any food allergies or dietary restrictions?” (Yes/No); “Are you keeping kosher for Passover?” (the study took place during the Passover holiday); and their demographics. There was no difference between conditions on any of these items and all results reported in the main text remain significant when controlling for them. Participants also read paragraphs of text and counted the number of occurrences of the letter “e” in each paragraph, examined Shepard-Metzler (1971) figures to determine whether sets of rotated figures were the same or not, completed a set of Raven’s Standard Progressive Matrices (Raven, 1983), and viewed 3 x 4 grids of numbers to identify pairs of numbers adding to 10 (Mazar, Amir, & Ariely 2008).

REFERENCES

Aaker, David A., and Donald E. Bruzzone (1985), "Causes of Irritation in Advertising," *Journal of Marketing*, 49 (2), 47–57.

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv (2005), "Loss Aversion Under Prospect Theory: A Parameter-free Measurement," *Management Science*, 53 (10), 1659–74.

Abelson, Reed (2011, November 16), "The Smokers' Surcharge," *The New York Times*, http://www.nytimes.com/2011/11/17/health/policy/smokers-penalized-with-health-insurance-premiums.html.

Aiken, Leona S. and Stephen G. West (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Ajzen, Icek (1991), "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes*, 50 (2), 179–211.

Ajzen, Icek and Martin Fishbein (1980), *Understanding Attitudes and Predicting Social Behavior*, Englewood Cliffs, NJ: Prentice-Hall.

Asch, Solomon E. (1951), "Effects of Group Pressure on the Modification and Distortion of Judgments," in *Group Leadership and Men*, ed. Harold Guetzkow, Pittsburgh: Carnegie Press, 177–90.

Barker, James R. (1992), "Tightening the Iron Cage: Concertive Control in Self-Managing Teams." *Administrative Science Quarterly*, 38 (3), 408–37.

Baumeister, Roy F., Arlene M. Stillwell, and Todd F. Heatherton (1994), "Guilt: An Interpersonal Approach," *Psychological Bulletin*, 115 (2), 243–67.

Bearden, William O. and Randall L. Rose (1990), "Attention to Social Comparison Information: An Individual Difference Factor Affecting Consumer Conformity," *Journal of Consumer Research*, 16 (4), 461–71.

Bem, Daryl J. (1972), "Self-Perception Theory," *Advances in Experimental Social Psychology*, 6, 1–62.

Berkowitz, Alan D (2004), "The Social Norms Approach: Theory, Research, and Annotated Bibliography." Retrieved January 19, 2018 from http://www.alanberkowitz.com/articles/social_norms.pdf.

Bettenhausen, Kenneth, and J. Keith Murnighan (1985), "The Emergence of Norms in Competitive Decision-Making Groups." *Administrative Science Quarterly,* 30 (3), 350–72.

Camerer, Colin F. (2005), "Three Cheers–Psychological, Theoretical, Empirical–for Loss Aversion," *Journal of Marketing Research*, 42 (2), 121–33.

Cialdini, Robert B. and Melanie R. Trost (1998), "Social Influence: Social Norms, Conformity and Compliance," in *The Handbook of Social Psychology*, Vol. 2, 4th ed., ed. Daniel T. Gilbert and Susan T. Fiske, York: McGraw-Hill, 151–92.

Cialdini, Robert. B., Raymond. R. Reno, and Carl. A. Kallgren (1990), "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places," *Journal of Personality and Social Psychology,* 58 (6), 1015–26.

Cialdini, Robert. B., Carl A. Kallgren, Raymond. R. Reno, (1991), "A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior," in *Advances in Experimental Social Psychology*, Vol. 24, ed. Leonard Berkowitz, San Diego: Academic Press, 201–34.

Cigna (2017), "Wellness Programs and Incentives," https://www.cigna.com/health-care-reform/wellness-and-incentives.

CMS (2013), "Overview: Final Rule for Health Insurance Market Reforms," https://www.cms.gov/CCIIO/Resources/Files/Downloads/market-rules-technical-summary-2-27-2013.pdf.

De Jong, Bart A., Katinka M. Bijlsma-Frankema, and Laura B. Cardinal (2014), "Stronger Than The Sum Of Its Parts? The Performance Implications Of Peer Control Combinations In Teams," *Organization Science*, 25 (6), 1703–21.

DellaVigna, Stefano and Devin Pope (2017), "What Motivates Effort? Evidence and Expert Forecasts," *Review of Economic Studies*, 85 (2), 1029–69.

Duke, Kristen and On Amir (2018), "Guilt Dynamics: Consequences of Temporally Separating Decisions and Actions," *Journal of Consumer Research*, forthcoming.

Ehrhart, Mark G., and Stefanie E. Naumann (2004), "Organizational Citizenship Behavior In Work Groups: A Group Norms Approach." *Journal of Applied Psychology,* 89 (6), 960–74.

Eriksson, Kimmo, Pontus Strimling, and Julie C. Coultas (2015), "Bidirectional Associations Between Descriptive And Injunctive Norms," *Organizational Behavior and Human Decision Processes,* 129, 59–69.

Evers, Ellen R. K., Yoel Inbar, Irene Blanken, and Linda D. Oosterwijk (2016), "When Do People Prefer Carrots to Sticks? A Robust 'Matching Effect' in Policy Evaluation," *Management Science,* 63 (12), 4261–76.

Fehr, Ernst and Urs Fischbacher (2004), "Social Norms and Human Cooperation," *Trends In Cognitive Sciences,* 8 (4), 185–90.

Fehr, Ernst and Ivo Schurtenberger (2018), "Normative Foundations of Human Cooperation," *Nature Human Behaviour,* 2, 458–68.

Fishbein, Martin, and Icek Ajzen (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research,* Reading, MA: Addison-Wesley.

Fishbein, Martin and Marco C. Yzer (2003), "Using Theory to Design Effective Health Behavior Interventions," *Communication Theory,* 13 (2), 164–83.

George, Jennifer M., and Gareth R. Jones (1997), "Organizational Spontaneity in Context," *Human Performance,* 10 (2), 153–70.

Gneezy, Uri and Aldo Rustichini (2000), "A Fine is a Price," *Journal of Legal Studies,* 29, 1–17.

Goldstein, Noah J., Robert B. Cialdini, and Vladas Griskevicius (2008), "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels," *Journal of Consumer Research,* 35 (3), 472–82.

Gonzalez-Mulé, Erik, David S. DeGeest, Brian W. McCormick, Jee Young Seong, and Kenneth G. Brown (2014), "Can We Get Some Cooperation Around Here? The Mediating Role of Group Norms on the Relationship Between Team Personality and Individual Helping Behaviors," *Journal of Applied Psychology,* 99 (5) 988–99.

Hackman, J. Richard. "Group Influences on Individuals in Organizations," in *Handbook of Industrial and Organizational Psychology,* eds. M.D. Dunnette & L.M. Hough, Palo Alto, CA: Consulting Psychologists Press, 199–267.

Haines, Michael, and Sherilynn F. Spear (1996), "Changing the Perception of the Norm: A Strategy to Decrease Binge Drinking Among College Students," *Journal of American College Health,* 45 (3), 134–40.

Hareli, Shlomo and Brian Parkinson (2008), "What's Social about Social Emotions?," *Journal for the Theory of Social Behaviour,* 38 (2), 131–56.

Hayes, Andrew F. (2013), *Introduction to Mediation, Moderation, and Conditional Process Analysis*, New York: Guilford Press.

Homonoff, Tatiana A. (2015), “Can Small Incentives Have Large Effects? The Impact of Taxes versus Bonuses on Disposable Bag Use,” working paper: Wagner Graduate School of Public Service, New York University.

Johnson, Eric J., and Daniel Goldstein (2003), "Do Defaults Save Lives?," *Science*, 302 (5649), 1338–39.

Kahneman, Daniel and Amos Tversky (1979), “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica*, 66 (3), 497–527.

Tversky, Amos, and Daniel Kahneman (1986), "Rational Choice and the Framing of Decisions., *Journal Of Business*, 59 (4), S251–S278.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1991), "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," *The Journal of Economic Perspectives*, 5 (1), 193–206.

Karmarkar, Uma R. and Bryan Bollinger (2015), "BYOB: How Bringing Your own Shopping Bags Leads to Treating Yourself and the Environment," *Journal of Marketing*, 79 (4), 1–15.

Keltner, Dacher and Cameron Anderson (2000), “Saving Face for Darwin: The Functions and Uses of Embarrassment,” *Current Directions in Psychological Science*, 9 (6), 187–92.

Keltner, Dacher and Brenda N. Buswell (1997), “Embarrassment: Its Distinct Form and Appeasement Functions,” *Psychological Bulletin*, 122 (3), 250–70.

Keltner, Dacher, Randall C. Young, and Brenda N. Buswell (1997), "Appeasement in Human Emotion, Social Practice, and Personality," *Aggressive Behavior*, 23 (5), 359–74.

kff (2016), "2016 Employer Health Benefits Survey," https://www.kff.org/report-section/ehbs-2016-summary-of-findings/.

Khan, Uzma and Ravi Dhar (2006), "Licensing Effect in Consumer Choice," *Journal of Marketing Research*, 43 (2) 259–66.

Krijnen, Job MT, David Tannenbaum, and Craig R. Fox (2017), "Choice Architecture 2.0: Behavioral Policy as an Implicit Social Interaction," *Behavioral Science & Policy*, 3 (2), i-18.

Lapinski, Maria K. and Rajiv N. Rimal (2005), "An Explication of Social Norms," *Communication Theory*, 15 (2), 127–47.

Larrick, Richard P., Jack B. Soll, and Ralph L. Keeney (2015), "Designing Better Energy Metrics for Consumers," *Behavioral Science and Policy*, 1 (1), 63–75.

Lennox, Richard D. and Raymond N. Wolfe (1984), "Revision of the Self-Monitoring Scale," *Journal of Personality and Social Psychology*, 46 (6), 1349–64.

Lewis, Helen B. (1971), "Shame and Guilt in Neurosis," *Psychoanalytic Review*, 58 (3), 419–38.

Marschall, Donna, Jennifer Sanftner, and June P. Tangney (1994), "The State Shame and Guilt Scale," *Fairfax, VA: George Mason University*.

McKenzie, Craig R. M. (2004), "Framing Effects in Inference Tasks–And Why They Are Normatively Defensible," *Memory & Cognition*, 32 (6), 874–85.

McKenzie, Craig R. M., Michael J. Liersch, and Stacey R. Finkelstein (2006), "Recommendations Implicit in Policy Defaults," *Psychological Science*, 17 (5), 414–20.

Miller, D. T., & Prentice, D. A. (1996), "The Construction of Social Norms and Standards," In *Social Psychology: Handbook of Basic Principles*, eds. E. T. Higgins & A. W. Kruglanski, New York, NY, US: Guilford Press, 799–829.

Moon, Junyean, Doren Chadee, and Surinder Tikoo (2008), "Culture, Product Type, and Price Influences on Consumer Purchase Intention to Buy Personalized Products Online," *Journal of Business Research*, 61 (1), 31–39.

Mulder, Laetitia B. (2008), "The Difference Between Punishments and Rewards in Fostering Moral Concerns in Social Decision Making," *Journal of Experimental Social Psychology*, 44 (6), 1436–43.

Mulder, Laetitia B. (2016), "When Sanctions Convey Moral Norms," *European Journal of Law and Economics*, 1–12.

Mullen, Brian, Jennifer L. Atkins, Debbie S. Champion, Cecelia Edwards, Dana Hardy, John E. Story, and Mary Vanderklok (1985), "The False Consensus Effect: A Meta-Analysis of 115 Hypothesis Tests," *Journal of Experimental Social Psychology*, 21 (3), 262–83.

Nasiry, Javad and Ioana Popescu (2011), "Dynamic Pricing with Loss-Averse Consumers

and Peak-end Anchoring," *Operations Research,* 59 (6), 1361–68.

National Conference of State Legislatures (2017), "State Plastic and Paper Bag Legislation: Fees, Taxes, and Bans," http://www.ncsl.org/research/environment-and-natural-resources/plastic-bag-legislation.aspx.

Neighbors, Clayton, Mary E. Larimer, and Melissa A. Lewis (2004), "Targeting Misperceptions of Descriptive Drinking Norms: Efficacy of a Computer-delivered Personalized Normative Feedback Intervention," *Journal of Consulting and Clinical Psychology,* 72 (3), 434–47.

Ouellette, J. A., & Wood, W. (1998), "Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior," *Psychological Bulletin,* 124 (1), 54–74.

Paluck, Elizabeth Levy, and Hana Shepherd (2012), "The Salience of Social Referents: A Field Experiment on Collective Norms and Harassment Behavior in a School Social Network," *Journal Of Personality And Social Psychology,* 103 (6), 899–915.

Payne, Rodger A. (2001), "Persuasion, Frames, and Norm Construction," *European Journal of International Relations,* 7 (1), 37–61.

Poortinga, Wouter (2017), "Results of a Field Experiment to Reduce Coffee Cup Waste. Summar Report to Bewley's Tea & Coffee UK Ltd," working paper: Welsh School of Architecture and School of Psychology, Cardiff University.

Poortinga, Wouter, Lorraine Whitmarsh, and Christine Suffolk (2013), "The Introduction of a Single-Use Carrier Bag Charge in Wales: Attitude Change and Behavioural Spillover Effects," *Journal of Environmental Psychology,* 36, 240–47.

Robinson, Eric, Jason Thomas, Paul Aveyard, and Suzanne Higgs (2014), "What Everyone Else is Eating: a Systematic Review and Meta-Analysis of the Effect of Informational Eating Norms on Eating Behavior," *Journal of the Academy of Nutrition and Dietetics,* 114 (3), 414–29.

Ross, Lee, David Greene, and Pamela House (1977), "The “False Consensus Effect”: An Egocentric Bias In Social Perception And Attribution Processes," *Journal Of Experimental Social Psychology,* 13 (3), 279–301.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius (2007), "The Constructive, Destructive, and Reconstructive Power of Social Norms," *Psychological Science,* 18 (5), 429–34.

Shah, Avni M., James R. Bettman, Peter A. Ubel, Punam Anand Keller, and Julie A. Edell (2014), "Surcharges Plus Unhealthy Labels Reduce Demand for Unhealthy Menu Items," *Journal of Marketing Research,* 51(6), 773–89.

Sher, Shlomi and Craig R. M. McKenzie (2006), "Information Leakage from Logically Equivalent Frames," *Cognition*, 101 (3), 467—94.

Smith-McLallen, Aaron, and Martin Fishbein (2008), "Predictors of Intentions to Perform Six Cancer-Related Behaviours: Roles for Injunctive and Descriptive Norms," *Psychology, Health and Medicine* 13 (4), 389–401.

Speck, Paul Surgi, and Michael T. Elliott (1997), "Predictors of Advertising Avoidance in Print and Broadcast Media," *Journal of Advertising*, 26 (3), 61–76.

Spiller, Stephen A., Gavan J. Fitzsimons, John G. Lynch Jr, and Gary H. McClelland (2013), "Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression," J*ournal of Marketing Research*, 50 (2), 277–88.

Starbucks Coffee Company (2017), "Cup and Materials," www.starbucks.com/responsibility/environment/cups-and-materials.

Stewart, Greg L., Stephen H. Courtright, and Murray R. Barrick (2012), "Peer-Based Control in Self-Managing Teams: Linking Rational and Normative Influence with Individual and Group Performance," *Journal of Applied Psychology*, 97 (2), 435–47.

Swaab, Roderick I., and Adam D. Galinsky (2015), "Egalitarianism Makes Organizations Stronger: Cross-National Variation in Institutional and Psychological Equality Predicts Talent Levels and the Performance of National Teams," *Organizational Behavior and Human Decision Processes,* 129 (2015), 80–92.

Tangney, June P., Rowland S. Miller, Laura Flicker, and Deborah H. Barlow (1996), "Are Shame, Guilt, and Embarrassment Distinct Emotions?," *Journal of Personality and Social Psychology*, 70 (6), 1256–69.

Tangney, June P. and Kurt W. Fischer (1995), "Self-Conscious Emotions: The Psychology of Shame, Guilt, Embarrassment, and Pride," *New York, NY*: Guilford Press.

Tankard, Margaret E. and Elizabeth Levy Paluck (2016), "Norm Perception as a Vehicle for Social Change," *Social Issues and Policy Review*, 10 (1), 181–211.

Tannenbaum, David, Chad J. Valasek, Eric D. Knowles, and Peter H. Ditto (2013), "Incentivizing Wellness in the Workplace: Sticks (Not Carrots) Send Stigmatizing Signals," *Psychological Science,* 24 (8), 1512–22.

Target (2017), "Corporate Responsibility: Sustainability," https://corporate.target.com/corporate-responsibility/sustainability/sustainable-products

Thaler, Richard H. (1980), “Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior & Organization*, 1, 39–60.

The Guardian (2018), “Starbucks trials 5p takeaway cup charge in attempt to cut waste”, https://www.theguardian.com/business/2018/feb/26/starbucks-5p-takeaway-cup-charge-london

Thomas, Gregory Owen, Wouter Poortinga, and Elena Sautkina (2016), "The Welsh Single-Use Carrier Bag Charge and Behavioural Spillover," *Journal of Environmental Psychology*, 47, 126–35.

Tracy, Jessica L., and Richard W. Robins (2004), "Putting the Self into Self-Conscious Emotions: A Theoretical Model," *Psychological Inquiry* 15 (2), 103–125.

White, Katherine M., Joanne K. Smith, Deborah J. Terry, Jaimi H. Greenslade, and Blake M. McKimmie (2009), "Social Influence in the Theory of Planned Behaviour: The Role of Descriptive, Injunctive, and Ingroup Norms," *British Journal of Social Psychology*, 48 (1), 135–58.